

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

Yao Bernard Fionko^{1,2*}, Serge Konan Yao^{1,2}, Raissa Diaby^{3,4}, Koffi Jean-Jacques Kouassi^{1,2}, Adama Simaga^{1,2}, Aderehime Haidara^{1,2}

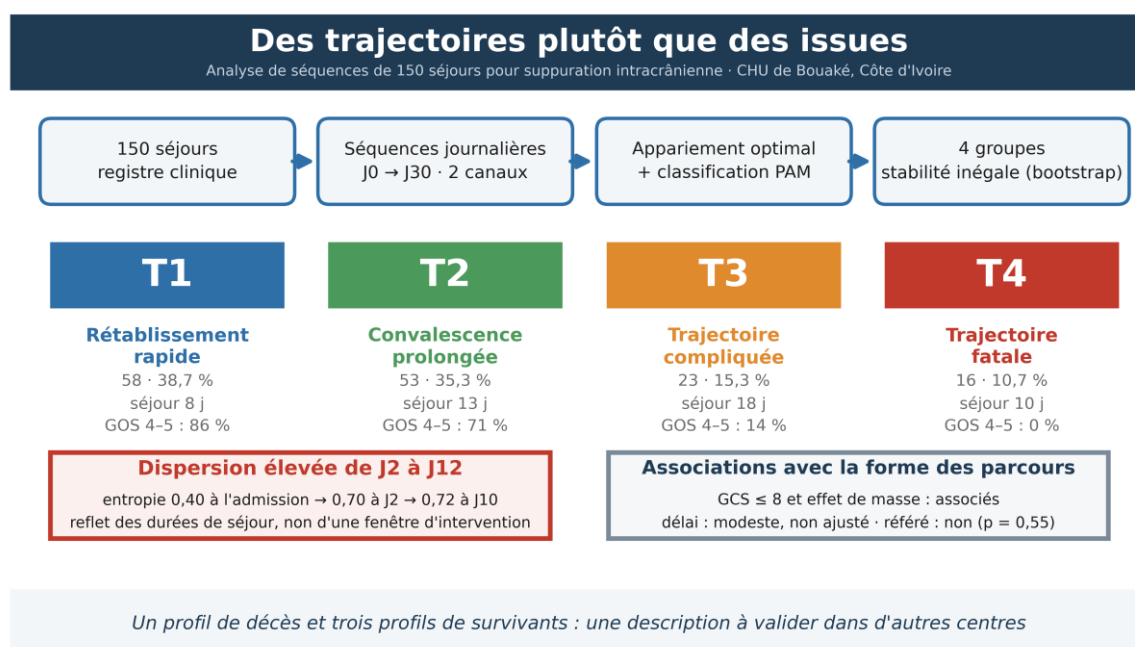
¹Department of Neurosurgery, Bouaké University Hospital, Bouaké, Côte d'Ivoire

²Faculty of Medical Sciences, Alassane Ouattara University, Bouaké, Côte d'Ivoire

³Department of Neurosurgery, Yopougon University Hospital, Abidjan, Côte d'Ivoire

⁴Faculty of Medical Sciences, Felix Houphouët Boigny University, Abidjan, Côte d'Ivoire

Graphical abstract



Graphical abstract. From the data source to the typology: clinical registry, sequence reconstruction, optimal matching and four trajectories, whose stability on bootstrap is uneven (one stable group, two patterns, one continuum).

ABSTRACT

Background. Studies of intracranial suppurative infections (ISI) in sub-Saharan Africa look at isolated variables — a delay, a Glasgow Coma Scale (GCS) score, a mortality rate — and rarely describe how a hospital stay unfolds. We asked whether a day-by-day reconstruction of the pathway tells us something that conventional prognostic analyses do not.

Methods. Single-centre cohort of 150 admissions (148 patients) for ISI at Bouaké University Hospital, Côte d'Ivoire, 2016–2026. Each stay became a 31-day sequence on two channels, care (seven states) and neurological severity (three GCS bands), using rules fixed before clustering. Sequences were compared by optimal matching and grouped by partitioning around medoids. Stability was tested by bootstrap and eleven sensitivity analyses, admission factors by discrepancy analysis and Firth regressions. The study is exploratory.

Results. The 150 admissions produced 100 distinct sequences, grouped into four profiles: rapid recovery (T1, 38.7%), prolonged convalescence (T2, 35.3%), complicated course (T3, 15.3%) and death (T4, 10.7%). Length of stay, GCS, complication and death build the sequences, so their differences between groups are expected and validate nothing. Among other variables, only lesion diameter clearly separated the groups ($\epsilon^2 = 0.185$). Symptom-to-admission delay had only a weak effect ($\epsilon^2 = 0.049$; unadjusted $p = 0.017$) that disappeared after adjustment, and referral status was linked to nothing ($p = 0.55$). Entropy formed a plateau from day 2

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

to day 12, peaking at 0.72 on day 10. On bootstrap, only T4 was stable (Jaccard 0.985; T1 0.741, T3 0.672, T2 0.599). Four of eleven sensitivity analyses departed clearly from the main solution.

Conclusion. Stays fall into three survivor profiles of uneven clarity and one profile of death. The registry dates neither complications nor critical care, so the sequences mainly describe length of stay, complications and outcome; they say nothing about the order of events or about when to intervene. Multicentre validation and a registry that dates complications and critical care are needed.

KEYWORDS: intracranial suppurative infection; brain abscess; subdural empyema; care trajectory; sequence analysis; optimal matching; care pathway; sub-Saharan Africa; implementation science; global neurosurgery

INTRODUCTION

Variables taken one at a time

Intracranial suppurative infections (brain abscess, subdural and extradural empyema, ventriculitis) remain common in low- and middle-income countries and are more often fatal than in high-resource series [1, 2]. The African literature is mostly descriptive: epidemiological profiles, portals of entry, bacteriology, prognostic factors [3-6]. It examines each variable separately against a binary outcome.

This is useful, and we have done it ourselves on this cohort with prognostic scores and quality indicators. But such analyses treat the patient as a list of variables and the hospital stay as a black box. Two patients admitted the same day with the same Glasgow score and the same collection do not necessarily live through the same stay. One is operated on the next day and is discharged on day 8. The other is operated on the next day, reoperated on day 6 for a re-accumulated collection, and is discharged on day 19. A conventional analysis nonetheless gives them the same value for "surgery" and for "survival".

Telling the stay rather than summarising it

Sequence analysis starts from a different idea. Each patient is described by the series of states he or she goes through, day after day. The distance between these series is then measured [7], and those that resemble each other are grouped. The method was born in sociology, to study careers and family trajectories [8, 9], and took shape in the 2000s [10-14]. It has since spread to health services research: pathways after a first hospitalisation for chronic obstructive pulmonary disease (COPD) [15], before a first diagnosis of schizophrenia [16], inequalities in pathways after a diagnosis of diabetes [17], medication use [18, 19], prenatal care pathways [20]. It sits alongside other approaches to pathways, such as process mining or group-based trajectory modelling [21, 22], which answer other questions.

For a neurosurgical department, the interest is to see the shape of the stay: the order of events, their dates, their durations [23]. One can then ask questions that conventional analyses do not ask. Do pathways separate at reception, in the operating theatre or later? Are there recurring shapes of stay? Do the factors present at admission weigh only on the outcome, or on the whole stay?

These questions have practical consequences. Knowing that a delay increases mortality says that it must be shortened; knowing when stays diverge would say where to intervene. This is why pathways are of growing interest to implementation science [24-26], which is concerned less with the efficacy of a treatment than with the concrete conditions in which it is delivered. Care systems are readily seen there as complex adaptive systems [27, 28], which argues for following pathways rather than isolated variables.

A method little used in African neurosurgery

We found no application of sequence analysis to intracranial suppurative infections, or to a neurosurgical department in sub-Saharan Africa. Two obstacles probably explain this. First, the method was designed for dense administrative data [29, 30] that most hospitals in sub-Saharan Africa do not have. Second, neurosurgery reasons in terms of risk factors and outcomes, rarely in terms of pathways. The first obstacle weighs less than one might think: a few dated landmarks from a well-kept clinical registry are enough to reconstruct a sequence. In return, granularity is lower and the reconstruction rules must be written down in black and white.

Objectives

We had four objectives: to reconstruct, from an ordinary registry, the daily course of each patient's stay; to describe the structure of these stays over time; to derive a typology from them and measure its stability; and to look for which factors known at admission are linked to the shape of the stay, and not only to its outcome.

Our hypothesis, outlined in Figure 1, is that the sequence of states is a property of the care system, which cannot be reduced to the sum of individual prognostic variables. This single-centre observational study can explore it, not demonstrate it.

Cadre conceptuel : la trajectoire comme unité d'analyse

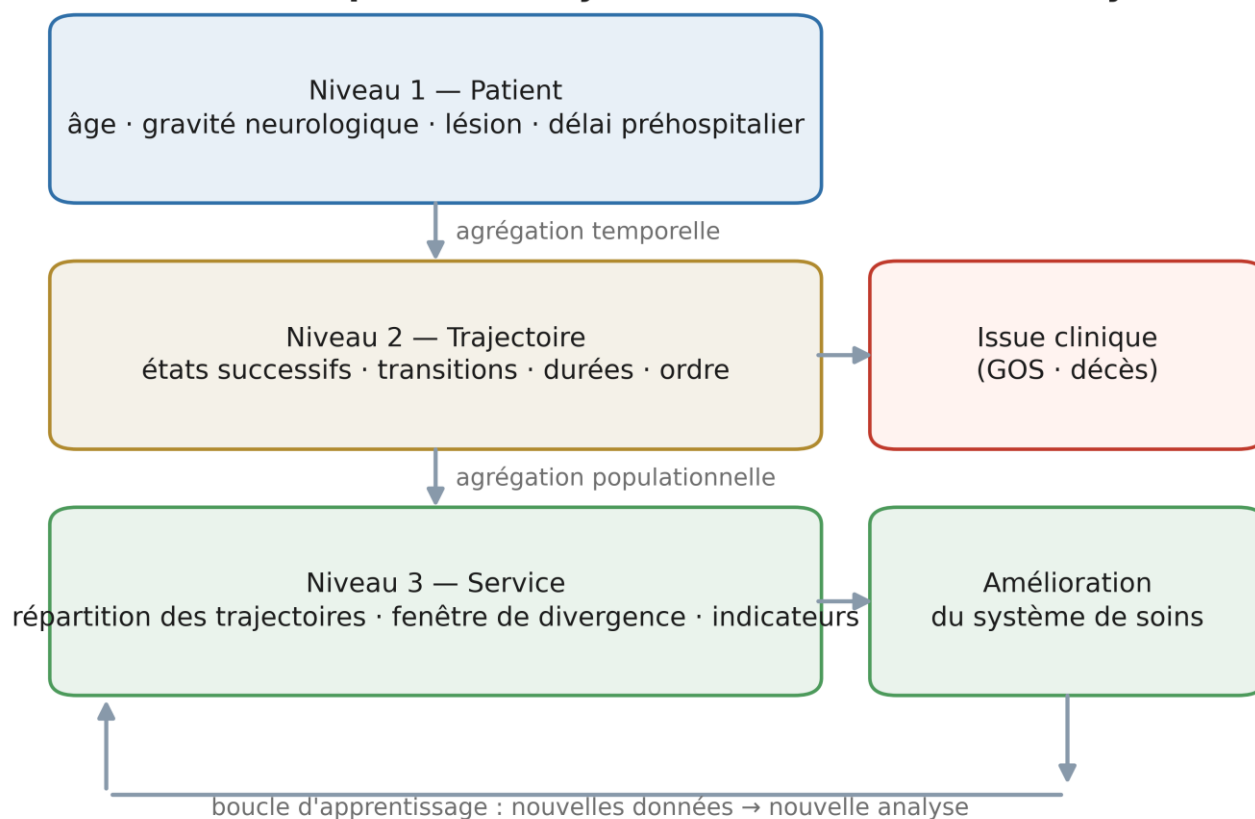


Figure 1. Conceptual framework: the trajectory as the unit of analysis. Three levels of analysis. Level 1: individual characteristics at admission, the object of conventional prognostic analyses. Level 2: the trajectory, the ordered series of clinical and organisational states a patient goes through. Level 3: the department, whose performance can be described by the distribution of trajectories. The bottom loop depicts a learning health system.

MATERIALS AND METHODS

Study type and setting

Observational cohort study conducted in the neurosurgery department of Bouaké University Hospital, the public referral hospital of central Côte d'Ivoire. Data collection was mixed: retrospective review of records, then prospective collection in 2025-2026. The department has a CT scanner, an operating theatre, a bacteriology laboratory and a general, non-specialised intensive care unit. The report follows STROBE [31] (checklist in S5) and, the data coming from a routine registry, the applicable items of RECORD. The study does not build a predictive model and therefore does not fall under TRIPOD; we used it only as an internal check. The flow of the study is outlined in Figure 2.

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

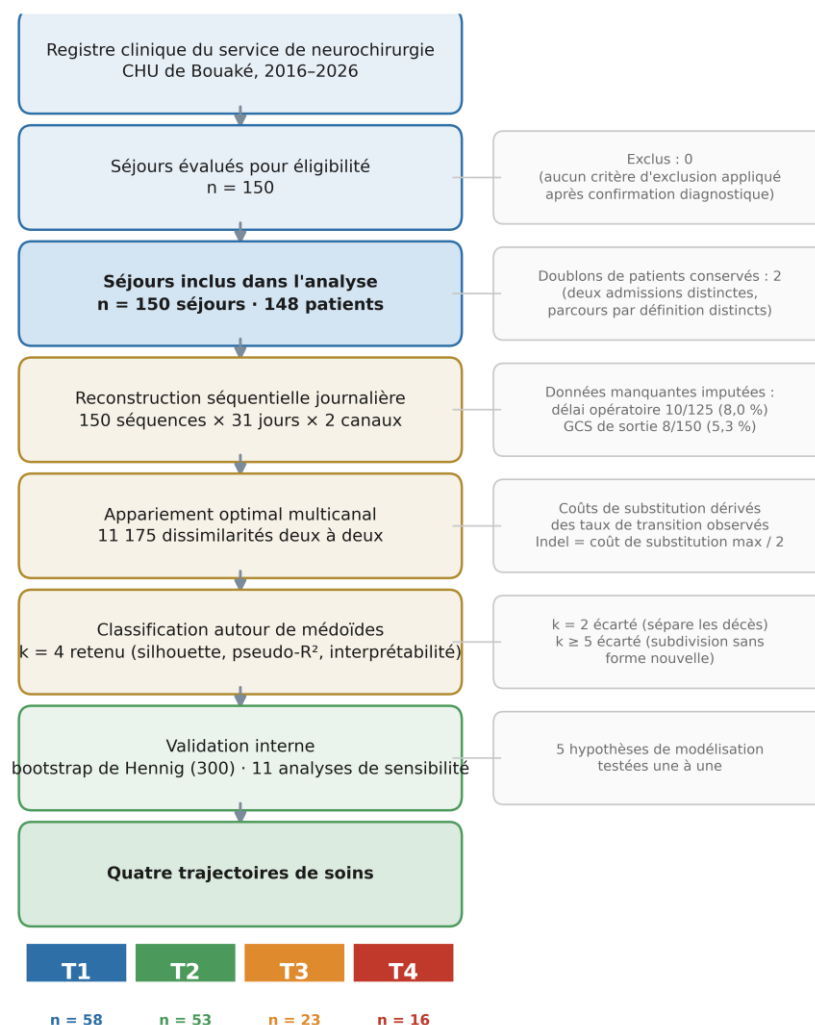


Figure 2. Study flow and analytical chain. From registry to typology: constitution of the cohort (top), sequence analysis chain (centre), validation and result (bottom). No stay was excluded after application of the inclusion criteria.

Population and sources

We included all stays for documented intracranial suppurative infection between 2016 and 2026, at any age: brain abscess, subdural or extradural empyema, ventriculitis, pre-suppurative cerebritis, scalp suppuration with cranial involvement. Diagnosis rested on a suggestive clinical picture and compatible cross-sectional imaging, with or without bacteriological proof. The definitions are those of the main article of the series (brain abscess: parenchymal collection with ring enhancement after injection). Cross-sectional imaging being required for inclusion, all stays had it: this figure therefore does not measure access to imaging. The unit of analysis is the stay. The cohort comprises 150 stays for 148 patients; two patients were admitted twice, and their two stays are distinct. The data come from a structured registry (12 sections, 112 variables) built from hospitalisation records, operative reports and consultation registers. To reconstruct the sequences, we used the date of admission, the symptom-to-admission delay, the admission and discharge GCS, the therapeutic strategy, the admission-to-intervention delay, the occurrence of a complication, reoperation or recurrence, the length of stay and the vital status. The Glasgow Outcome Scale (GOS) was not collected separately: it was assigned from sequelae and autonomy, according to a correspondence rule fixed in advance, and could not be assigned in 8 survivors. It is therefore not a measure of outcome independent of the discharge GCS and sequelae.

Construction of state sequences

Each stay becomes a sequence of 31 positions (days 0 to 30 from admission) on two parallel channels. The care channel has seven exclusive states: awaiting surgery, day of surgery, simple postoperative course, complicated postoperative course, medical treatment only, discharged alive and death. The last two are absorbing: once reached, the patient stays there until day 30, which makes it possible to compare stays of different durations. The severity channel has three ordered bands ($GCS \geq 13$, $9-12$, ≤ 8) and the two absorbing states.

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

As the GCS is recorded only at admission and at discharge, this channel is a step function: the admission value up to a pivot date, the discharge value thereafter. The pivot date is the day after the intervention in operated patients, and the middle of the stay in medically treated patients.

In all, a sequence summarises only seven pieces of information from the registry: surgery and operative delay, complication or reoperation, length of stay, death, admission GCS and discharge GCS. The registry does not date complications, which we place by convention from the third postoperative day, and says nothing about intensive care. The order of states is therefore imposed by the construction (waiting, day of surgery, simple or complicated postoperative course, discharge or death); only their dates and durations vary.

This reconstruction rests on five assumptions (Table 1). None comes directly from the data; we fixed them in advance and documented them. Four were subjected to a sensitivity analysis. The GCS pivot date was not: analysis AS3 removes the whole severity channel, which is a different question. The two imputations of missing data were not tested either. The dictionary of states, the rules, an annotated example and the code are in S2, S3 and S6.

Table 1. Modelling assumptions and missing data: declaration and assessment

Modeling assumptions and missing data: declaration and assessment					
Assumption or Extent		Rule applied		Sensitivity analysis	Influence (ARI)
Observation horizon		–	31 days (D0-D30), longer than the maximum length of stay (26 days)	21-day horizon (AS6)	Negligible (0.992)
Absorbing states		–	Discharge and death occupied until the end of the horizon	Reduced horizon (AS6)	Negligible (0.992)
Missing delay	operative 10/125 (8.0%)	operated	Imputation to the median of operated patients (2 days)	Exclusion of these 10 stays (AS7)	Low (0.908)
Date of occurrence of complications		36 stays concerned	Complicated state assigned from postoperative D+3	D+2 and D+5 variants (AS1, AS2)	Moderate at D+2 (0.534) · low at D+5 (0.893)
GCS interpolation (pivot date)	150 stays		Step function: admission GCS then discharge GCS, pivot at postoperative D+1 or mid-stay	No dedicated analysis (AS3 removes the whole severity channel)	Not assessed (AS3: 0.265)
Missing diameter	lesion 6/150 (4.0 %)		Imputation to the median, Firth model only	–	–
Missing GCS	discharge 8/150 (5.3 %)		Admission value carried forward	–	–

None of these assumptions follows from the observed data: each was fixed in advance and documented in S2. Four were assessed by a sensitivity analysis (AS1 to AS11; see Results, Sensitivity analyses). ARI: adjusted Rand index relative to the main partition.

Dissimilarities, clustering and validation

The distance between two sequences is computed by optimal matching: it is the minimum cost of the insertions, deletions and substitutions needed to transform one into the other [7, 8]. In the care channel, the substitution cost is derived from the observed transition rates [32, 33]: two states between which patients often move are judged close. In the severity channel, it is proportional to the gap between Glasgow bands. The insertion-deletion cost equals half the maximum substitution cost, and the two channels are added with equal weights. The formalisation, the parameterisation and the comparison of metrics are in S6 and S8. As optimal matching has been criticised [34, 35], several sensitivity analyses concern the metric.

We clustered the sequences with the partitioning around medoids (PAM) algorithm [36], which represents each group by a real patient rather than by a virtual centre. The number of groups was examined from 2 to 8 on the basis of the mean silhouette [37], the pseudo-R² and clinical readability. The silhouette is maximal at k = 2, but this solution merely sets the deaths apart, which is known without sequence analysis.

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

We checked internal validity at four levels: the metric (AS4, AS10, AS11), the clustering algorithm (AS8, AS9), the groups themselves (Hennig bootstrap [38], 300 resamples, Jaccard index, all original patients being reassigned to the medoids of each sample) and the associations (bootstrap CIs of the pseudo- R^2 , profile likelihood of the odds ratios). A group is considered stable above 0.75, a real pattern between 0.60 and 0.75, and non-reproducible below 0.50. These thresholds do not qualify the 0.50-0.60 interval; we read it as doubtful.

Eleven sensitivity analyses (AS1 to AS11), planned in the protocol (S1), redid the whole chain: reconstruction, distances, clustering. Each is compared with the main solution by the adjusted Rand index [39].

Statistical analyses

Continuous variables are given as median and interquartile range (IQR), with the bootstrap 95% CI of the median. Categorical variables are given as counts and percentages, with the Wilson CI [40]. Groups are compared with the Kruskal-Wallis and chi-square tests, with an effect size: epsilon squared (ϵ^2) [41] for continuous variables, Cramér's V for categorical ones. Our ϵ^2 is the bias-corrected estimator, $(H - k + 1)/(n - k)$, which can take small negative values.

Discrepancy analysis [42] measures the association between a covariate and the set of sequences. It provides a pseudo- R^2 , the share of sequence variability explained by the covariate, and a pseudo-F statistic whose null distribution is obtained by 1,000 permutations. It does not depend on any choice of k . Membership of each group was modelled by one-versus-all logistic regression with Firth penalisation [43, 44], made necessary by complete separation: no T1 patient had a GCS ≤ 8 . The models included age, symptom-to-admission delay, GCS ≤ 8 , referral status, maximal lesion diameter and mass effect, all known at admission or before (Figure 3). This diagram is not a causal graph: we estimate no causal effect. The smallest group has only sixteen events, so the estimates are exploratory. Finally, the admission GCS feeds the severity channel: its link with the shape of the sequences is partly mechanical.

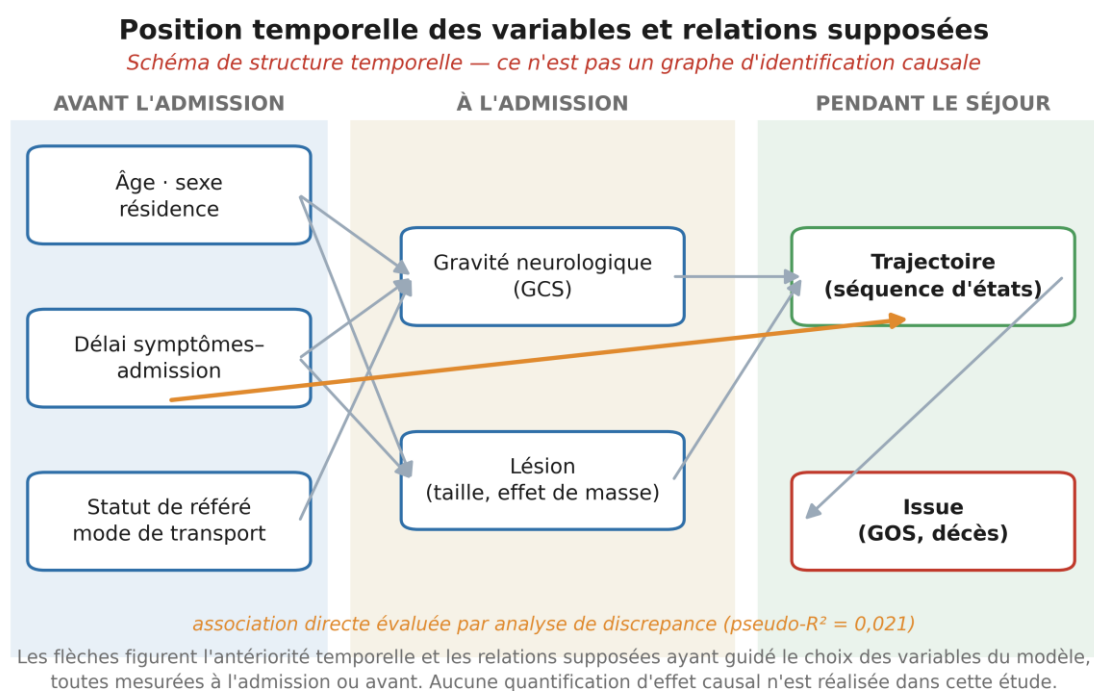


Figure 3. Temporal position of variables and assumed relationships. Temporal order of the variables and assumed relationships that guided their choice in the models: all are measured at admission or before. The orange arrow depicts the direct association between prehospital delay and the shape of the trajectory, assessed by discrepancy analysis. This is not a causal graph: no causal effect is estimated here and no adjustment is justified by this diagram.

To measure the diversity of stays on each day, we used Shannon entropy, normalised by the logarithm of the number of states. It equals 0 if all patients are in the same state and 1 if they are spread equally. It measures dispersion and not severity: a rise means that stays are diversifying, not that they are worsening. As discharges and deaths accumulate over time, entropy depends directly on the distribution of lengths of stay.

We did not correct for multiple comparisons, the study being exploratory: the p values are descriptive. Among the ten covariates of the discrepancy analysis, only $GCS \leq 8$ and mass effect ($p = 0.001$) would pass the Bonferroni threshold of 0.005. Missing data were rare (Table 1): operative delay in 10 of the 125 operated patients (8.0%), lesion diameter in 6 (4.0%), discharge GCS in 8 (5.3%; discharges against medical advice or lost to follow-up before assessment). No imputation was made for the descriptive

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

variables. Analyses were run in Python 3.12; optimal matching, clustering, discrepancy and Firth regression were programmed directly (code in S3).

Why this method

Table 2 compares the possible methods. Logistic regression answers "who will do badly?", survival models "when will the event occur?". Multistate models describe transitions [45], but require that they all be observed and dated, which our registry does not allow. Process mining reconstructs the hospital course [46, 47], but assumes time-stamped data from an information system. Sequence analysis answers a different question (how is the stay organised over time?) at the price of reconstruction assumptions that must be declared. These methods complement each other. Other approaches, such as mixture hidden Markov models [48] or mixtures of exponential-distance models [49], were not explored.

Table 2. Why sequence analysis: comparison with alternative approaches

Approach	Question it answers	Data requirement	Why it was not retained here
Logistic regression	Who will do badly?	Admission variables and outcome	Ignores the order and date of events; already applied to this cohort
Survival models (Cox)	When does the event occur?	Date of event and censoring	Ignores intermediate states and the shape of the pathway
Multistate models	How do transitions occur?	All transitions observed and dated	The registry dates neither complications nor the move to critical care
Process mining	How does the hospital process unfold?	Exhaustive time-stamped event log	No hospital information system
Group-based trajectory modelling	How does a continuous variable evolve?	Repeated measures of a continuous variable	No repeated measure available (GCS recorded twice)
State sequence analysis	How is the pathway organised over time?	Dated landmarks allowing daily reconstruction	Approach retained; requires declaring the reconstruction rules

This table justifies the choice of method without ranking the approaches, which answer different questions. Process mining and multistate models would have been preferable if the registry had time-stamped all transitions.

Ethics and patient participation

The study respects data confidentiality, after anonymisation, and the principles of the Declaration of Helsinki. It was approved internally by the neurosurgery department of Bouaké University Hospital. Informed consent was obtained for prospective inclusions (2025-2026); a waiver was granted for the retrospective phase, the data being anonymised.

Patients and the public did not take part in the design, conduct or analysis. Some questions (acceptability of delays, the experience of waiting before surgery, the experience of a reoperation) call for a qualitative approach with patients, which we plan next.

RESULTS

Description of the cohort

We analysed 150 stays, corresponding to 148 patients. Median age was 26 years (IQR 14-50), with two peaks: children and adolescents on one side, middle-aged adults on the other. The median delay between symptom onset and admission was 19 days (IQR 15-25; range 5-44). Seventy-seven stays (51.3%) concerned patients referred by another facility. An ambulance had been used for 40 stays (26.7%); the detailed mode of transport was known for 140 stays.

Median admission GCS was 13 (IQR 10-15) and 19 stays (12.7%) had a GCS $\leq$ 8. Cross-sectional imaging, a condition of inclusion, was available in all stays: CT alone in 101 cases, CT and MRI in 34, MRI alone in 13, transfontanellar ultrasound supplemented by CT in 2. This figure comes from the inclusion criterion; the delay in access to imaging is not documented. Median maximal lesion diameter was 36 mm (IQR 30-44).

One hundred and twenty-five stays (83.3%) involved surgery: 96 craniotomies and 11 external ventricular drains. Twenty-five (16.7%) were treated medically. Median operative delay was 2 days (IQR 1-3); 87 of the 125 operated patients (69.6%) were operated on within 48 hours. A pus sample was taken in 116 cases, with a positive culture in 50 (43.1%). Median length of stay was

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

11 days (IQR 8-15), and median duration of antibiotic therapy 7 weeks (IQR 6-7). A postoperative complication, reoperation or recurrence affected 36 stays (24.0%). Sixteen patients died (10.7%; 95% CI 6.7-16.6). Of 142 stays with an assessable GOS, 85 (59.9%) had a GOS of 4 or 5.

Highly varied stays

The 150 stays produced 100 distinct sequences; the most frequent was shared by only five patients (Supplementary Table S2). This number depends on the fineness of the coding (one day) and on the combination of operative date, length of stay and Glasgow bands. It says nothing about the standardisation of care. Daily transitions (Table 3, Figure 4) are almost deterministic. While awaiting surgery, the probability of remaining in waiting the next day was 0.562 and that of moving to the operating theatre 0.438. The day of surgery was always followed by a postoperative state. From the simple postoperative state, four outcomes were possible the next day: remain (0.860), be discharged (0.086), move to a complicated postoperative state (0.040) or die (0.013). The complicated state was very persistent (0.912), with a daily probability of discharge of 0.083. The matrix contains no backward transition: the order of states is imposed by the construction, and only their dates and durations vary.

Table 3. Matrix of daily transition probabilities between care states (whole cohort)

State on day t	Awaiting surgery	Day surgery	of Simple postop.	Complicated postop.	Medical only	Discharge	Death
Awaiting surgery	0.562	0.438	0.000	0.000	0.000	0.000	0.000
Day surgery	0.000	0.000	1.000	0.000	0.000	0.000	0.000
Simple postoperative	0.000	0.000	0.860	0.040	0.000	0.086	0.013
Complicated postoperative	0.000	0.000	0.000	0.912	0.000	0.083	0.005
Medical treatment only	0.000	0.000	0.000	0.000	0.896	0.096	0.008
Discharged alive	0.000	0.000	0.000	0.000	0.000	1.000	0.000
Death	0.000	0.000	0.000	0.000	0.000	0.000	1.000

Each row gives the distribution of states on day t+1 according to the state on day t; it sums to 1.000. Probabilities above 0.80 are in bold. Discharged alive and death are absorbing.

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

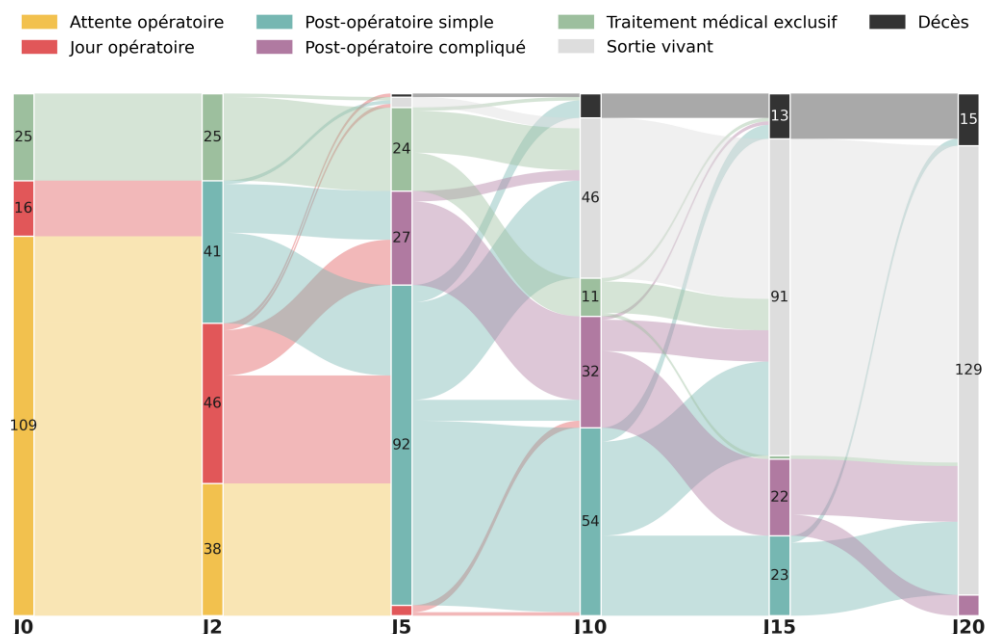


Figure 4. Flow diagram of care states at six time points of the stay. Each column shows the distribution of the cohort among the seven states on the indicated day; ribbons depict transitions between two time points. Counts are shown in blocks large enough. Complements Table 3.

Entropy (Figure 5) was 0.395 at admission and rose to 0.701 by the second day, when operative dates spread out. It fell back to 0.561 on days 5 and 6, when most patients are in the simple postoperative state, then rose again to a maximum of 0.717 on day 10. The two peaks differ by only 0.016. We have no confidence interval for the position of the maximum; the curve mainly forms a high plateau between D2 and D12.

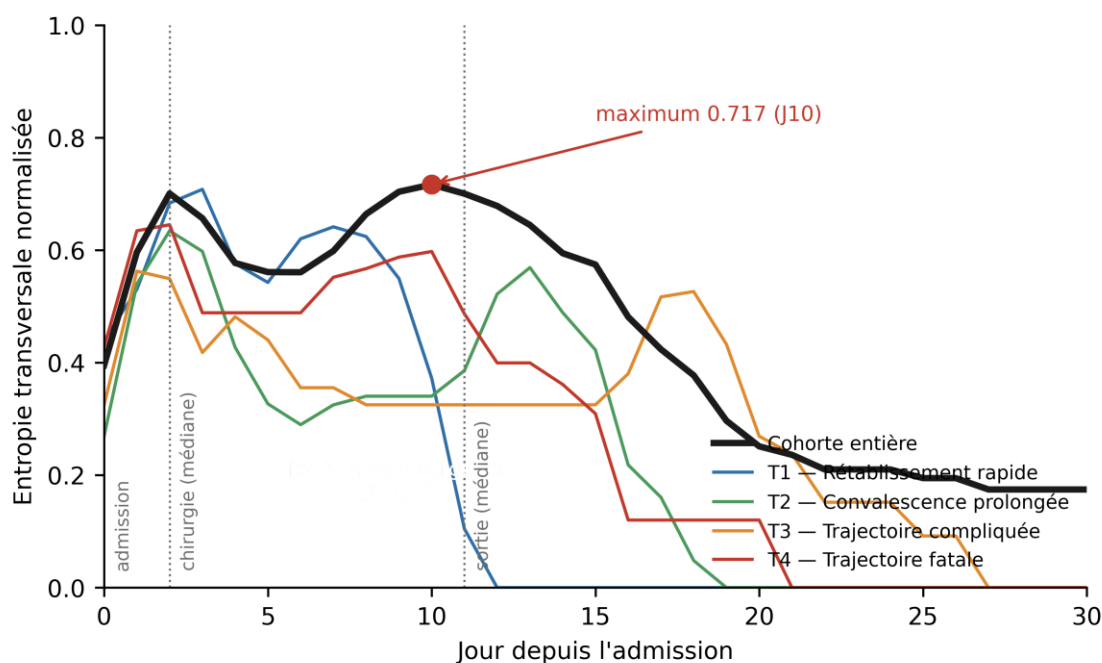


Figure 5. Normalised cross-sectional entropy of care states and position of the maximum. Entropy measures the diversity of states occupied at the same time by the cohort: 0 if all patients are in the same state, 1 if they are spread equally among the seven states. It does not measure severity: a rise means that stays are diversifying, not that they are worsening. Black curve: whole cohort; coloured curves: the four trajectories. The maximum (0.717) is reached on day 10; entropy is already 0.701 on day 2, the curve forming a high plateau from D2 to D12.

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

Stays are thus least dispersed at reception, where almost all patients are waiting for the theatre or receiving medical treatment. They disperse as soon as operations spread out and remain so until about the twelfth day, because patients who have been discharged, others in a simple postoperative course, others in a complicated course and others who have died are then side by side. As discharges feed entropy, its maximum falls close to the median length of stay (11 days). This profile therefore mainly describes the dispersion of lengths of stay, and it would be surprising if it were otherwise.

Nothing makes it possible to designate a window in which an organisational action would change the distribution of stays: an intervention study would be needed for that.

Four groups

Table 4 and Figure 6 detail the choice of the number of groups. The mean silhouette was 0.658 at $k = 2$, 0.435 at $k = 3$, 0.378 at $k = 4$, 0.406 at $k = 5$, 0.442 at $k = 6$, 0.472 at $k = 7$ and 0.481 at $k = 8$. The pseudo- R^2 , which increases mechanically with k , went from 0.533 to 0.896. At $k = 4$, the silhouette is the lowest from $k = 3$ to $k = 8$ and reflects a weak structure (< 0.5): we retained four groups because they are clinically readable, not because a statistical criterion requires it. The two-group solution separates the 16 deaths from the 134 survivors without teaching anything about trajectories. With four groups, the partition explains 77.8% (bootstrap 95% CI 73.3–82.9) of the variability of the dissimilarities. At five and six groups, T1 and T2 are redistributed into three or four subgroups, with no new clinical profile.

Table 4. Choice of the number of groups: partition quality criteria

k	ASW	Pseudo-R^2	Group sizes	Comment
2	0.658	0.533	134 / 16	Separates the deaths; no trajectory content
3	0.435	0.704	91 / 43 / 16	Does not distinguish the complicated trajectory
4 (retained)	0.378	0.778 [0.733–0.829]	58 / 53 / 23 / 16	Choice based on interpretability; lowest silhouette from $k = 3$ to $k = 8$
5	0.406	0.827	63 / 29 / 23 / 19 / 16	Redistributes T1 and T2 into three subgroups with no new shape
6	0.442	0.862	40 / 33 / 22 / 21 / 18 / 16	Further subdivision of T1 and T2
7	0.472	0.888	39 / 24 / 22 / 21 / 16 / 15 / 13	Small groups
8	0.481	0.896	36 / 23 / 23 / 16 / 16 / 15 / 13 / 8	Over-segmentation

ASW: mean silhouette width (optimal matching dissimilarity matrix). Pseudo- R^2 : share of the variability of the dissimilarities explained by the partition. At $k = 2$, the best silhouette corresponds to setting apart the sixteen deaths, without trajectory information. Retained solution: $k = 4$, whose silhouette (0.378) indicates a weak structure; the choice rests on clinical readability.

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

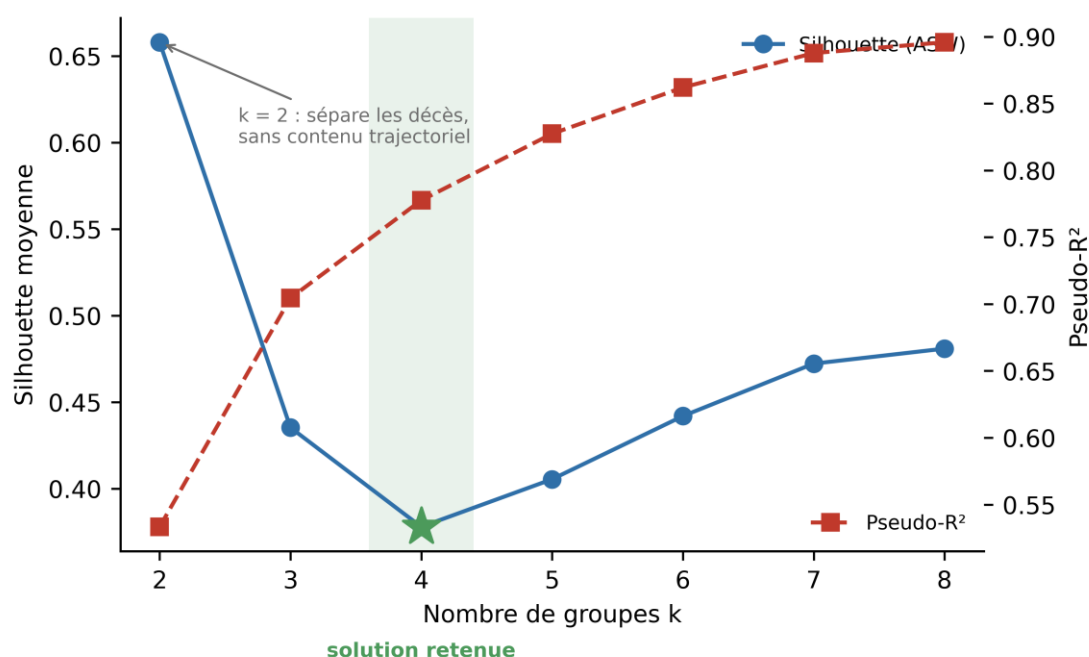


Figure 6. Criteria for choosing the number of groups. Mean silhouette (left scale) and pseudo-R² (right scale) according to the number of groups. The star marks the retained solution. The maximal silhouette at k = 2 corresponds to setting apart the sixteen deaths, a solution discarded for lack of trajectory content.

Figure 7 shows the stays one by one, Figure 8 their distribution day by day, Figure 9 the medoid sequences and Figure 10 the clinical synthesis. The characteristics of the groups are in Table 5, the mean time spent in each state in Table 6.

Table 5. Characteristics of the four care trajectories with effect sizes (n = 150 stays)

Variable	T1 Rapid recovery (n = 58)	T2 Prolonged convalescence (n = 53)	T3 Complicated trajectory (n = 23)	T4 Fatal trajectory (n = 16)	Total (n = 150)	Effect size	p value
Age (years)	30 (15–50)	23 (13–50)	27 (16–46)	16 (12–55)	26 (14–50)	$\epsilon^2 = -0.019$	0.97
Symptom-to-admission delay (days)	16 (13–22)	20 (16–25)	19 (17–22)	22 (18–30)	19 (15–25)	$\epsilon^2 = 0.049$	0.017
Admission GCS †	14 (13–15)	13 (10–15)	8 (6–10)	12 (8–13)	13 (10–15)	$\epsilon^2 = 0.323$	< 0.001
Admission GCS ≤ 8 †	0 (0.0 [0.0–6.2])	2 (3.8 [1.0–12.8])	12 (52.2 [32.7–71.0])	5 (31.3 [14.2–55.6])	19 (12.7)	V = 0.575	< 0.001
Maximal lesion diameter (mm)	31 (26–36)	38 (31–46)	43 (38–54)	40 (33–49)	36 (30–44)	$\epsilon^2 = 0.185$	< 0.001
Referred patient	29 (50.0 [37.5–62.5])	28 (52.8 [39.7–65.6])	11 (47.8 [29.2–67.0])	9 (56.2 [33.2–76.9])	77 (51.3)	V = 0.049	0.95
Transport by ambulance	9 (15.5 [8.4–26.9])	16 (30.2 [19.5–43.5])	10 (43.5 [25.6–63.2])	5 (31.3 [14.2–55.6])	40 (26.7)	V = 0.224	0.057
Surgical procedure †	39 (67.2 [54.4–77.9])	50 (94.3 [84.6–98.1])	22 (95.7 [79.0–99.2])	14 (87.5 [64.0–96.5])	125 (83.3)	V = 0.348	< 0.001
Operative delay (days) †	2 (2–3)	2 (1–3)	2 (1–2)	2 (1–2)	2 (1–3)	$\epsilon^2 = 0.039$	0.051

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

Variable	T1 Rapid recovery (n = 58)	T2 Prolonged convalescen ce (n = 53)	T3 Complicate d trajectory (n = 23)	T4 Fatal trajectory (n = 16)	Total (n = 150)	Effect size	p value
Complication, reoperation or recurrence †	7 (12.1 [6.0– 22.9])	9 (17.0 [9.2– 29.2])	18 (78.3 [58.1–90.3])	2 (12.5 [3.5– 36.0])	36 (24.0)	V = 0.543	< 0.001
Duration of antibiotic therapy (weeks)	7 (6–8)	7 (6–7)	6 (6–7)	6 (6–7)	7 (6–7)	$\epsilon^2 = 0.023$	0.09
Length of stay (days) †	8 (7–9)	13 (12–15)	18 (16–19)	10 (8–13)	11 (8–15)	$\epsilon^2 = 0.746$	< 0.001
Discharge GCS †	15 (15–15)	15 (12–15)	10 (8–12)	–	15 (11–15)	$\epsilon^2 = 0.566$	< 0.001
GOS (assessable stays) ‡	4-5 47/55 (85.5 [73.8–92.4])	35/49 (71.4 [57.6–82.2])	3/22 (13.6 [4.7–33.3])	0/16 (0.0 [0.0–19.4])	85/142 (59.9)	V = 0.656	< 0.001
Death †	0 (0.0 [0.0– 6.2])	0 (0.0 [0.0– 6.8])	0 (0.0 [0.0– 14.3])	16 (100.0 [80.6– 100.0])	16 (10.7)	V = 1.000	< 0.001

Continuous variables: median (IQR). Binary variables: count (percentage [Wilson 95% CI]). Effect sizes: epsilon squared (ϵ^2 , bias-corrected estimator; slight negative values are possible) for continuous variables; Cramér's V (2×4 contingency tables) for categorical variables. p: Kruskal-Wallis or chi-square, without correction for multiple comparisons (descriptive values). Operative delay: 125 operated stays. GOS: 142 assessable stays (T1 55, T2 49, T3 22, T4 16). Ambulance: recorded for all 150 stays (detailed mode of transport available for 140). † Variable used to build the sequences or determined by them: the difference between groups is expected and does not validate the typology. ‡ The GOS is assigned from sequelae and autonomy.

Table 6. Mean time spent in each state over the 31-day horizon (D0-D30), by trajectory

State	T1 (n = 58)	T2 (n = 53)	T3 (n = 23)	T4 (n = 16)
Awaiting surgery	1.53 (4.9 %)	1.96 (6.3 %)	1.52 (4.9 %)	1.31 (4.2 %)
Day of surgery	0.67 (2.2 %)	0.94 (3.0 %)	0.96 (3.1 %)	0.88 (2.8 %)
Simple postoperative	3.16 (10.2 %)	9.40 (30.3 %)	4.61 (14.9 %)	6.50 (21.0 %)
Complicated postoperative	0.64 (2.1 %)	1.55 (5.0 %)	11.57 (37.3 %)	1.62 (5.2 %)
Medical treatment only	2.83 (9.1 %)	0.77 (2.5 %)	0.70 (2.3 %)	1.19 (3.8 %)
Discharged alive	22.17 (71.5 %)	16.38 (52.8 %)	11.65 (37.6 %)	0.00 (0.0 %)
Death	0.00 (0.0 %)	0.00 (0.0 %)	0.00 (0.0 %)	19.50 (62.9 %)

Days (percentage of the 31-day horizon, D0-D30). Discharged alive and death are absorbing: the time attributed to them is the remainder of the horizon after the end of the stay, hence an indirect measure of the brevity of hospitalisation.

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

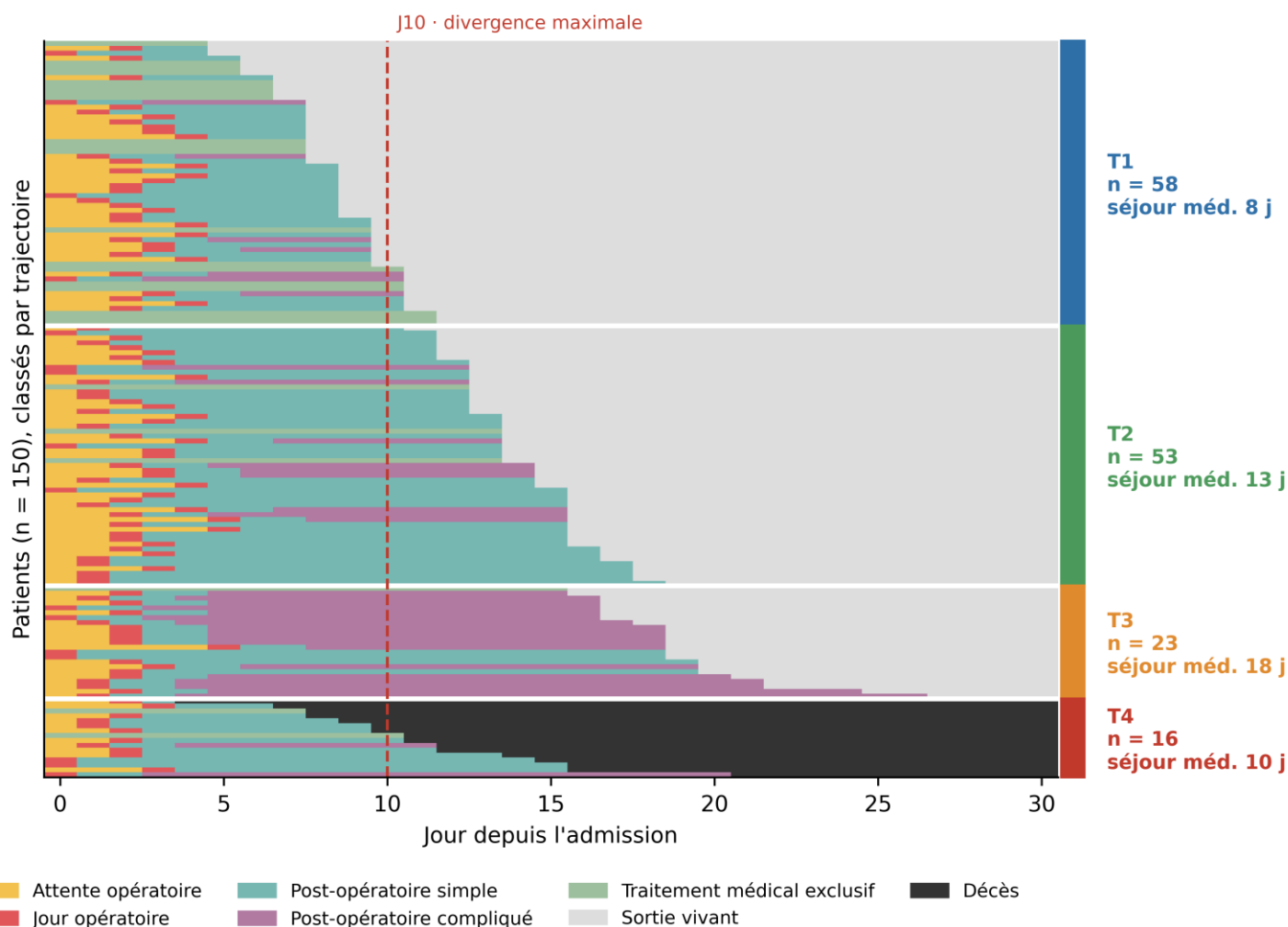


Figure 7. Index plot of the 150 care sequences, ordered by trajectory. Each row is one patient's stay over 31 days from admission; each colour is a care state. Stays are grouped by trajectory then ordered by increasing length. One can see at once the date of the operation, the length of convalescence, the occurrence of a complication and the outcome. The vertical line marks the tenth day.

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

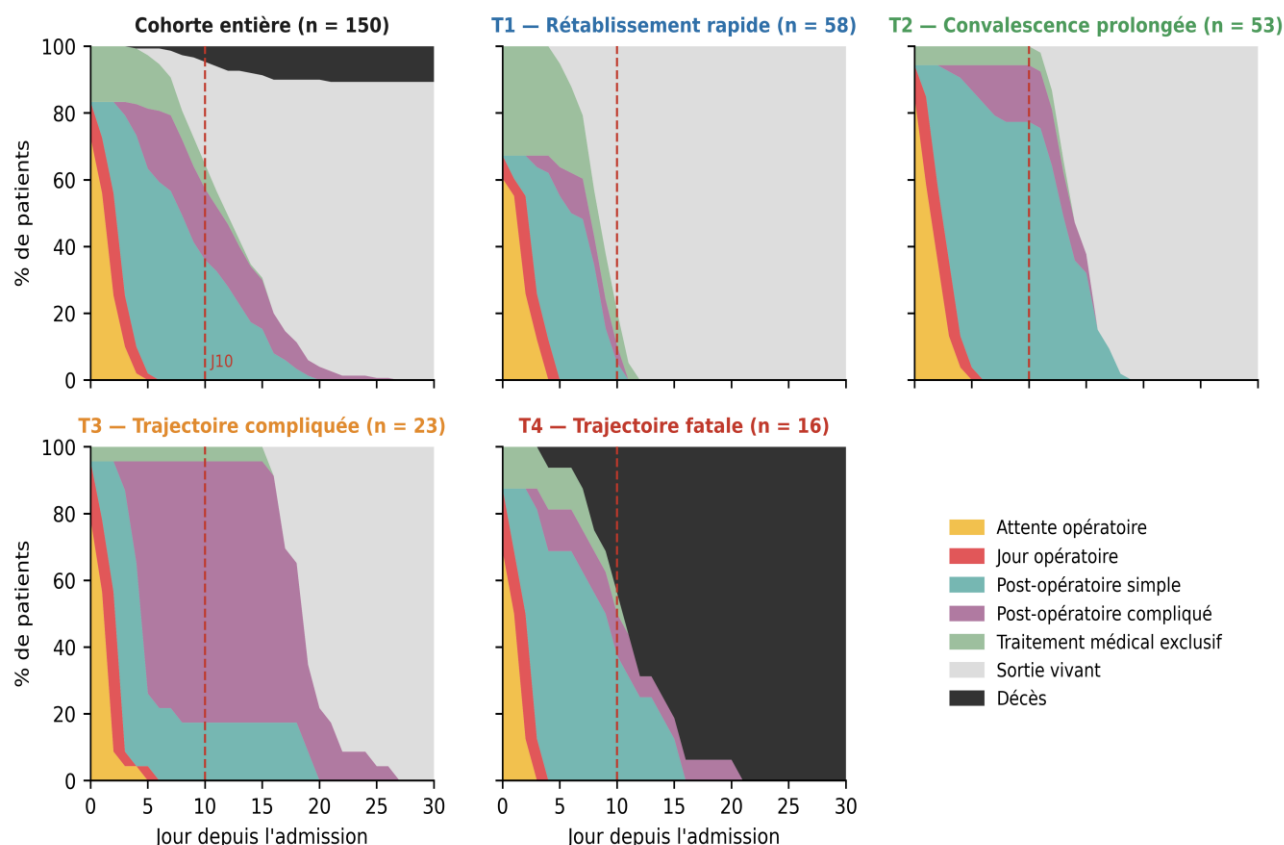


Figure 8. State distribution chronograms, whole cohort and by trajectory. For each day since admission, proportion of patients in each of the seven care states. T1: compressed convalescence. T2: stretched convalescence. T3: dominant complicated state. T4: progression of the absorbing state of death.

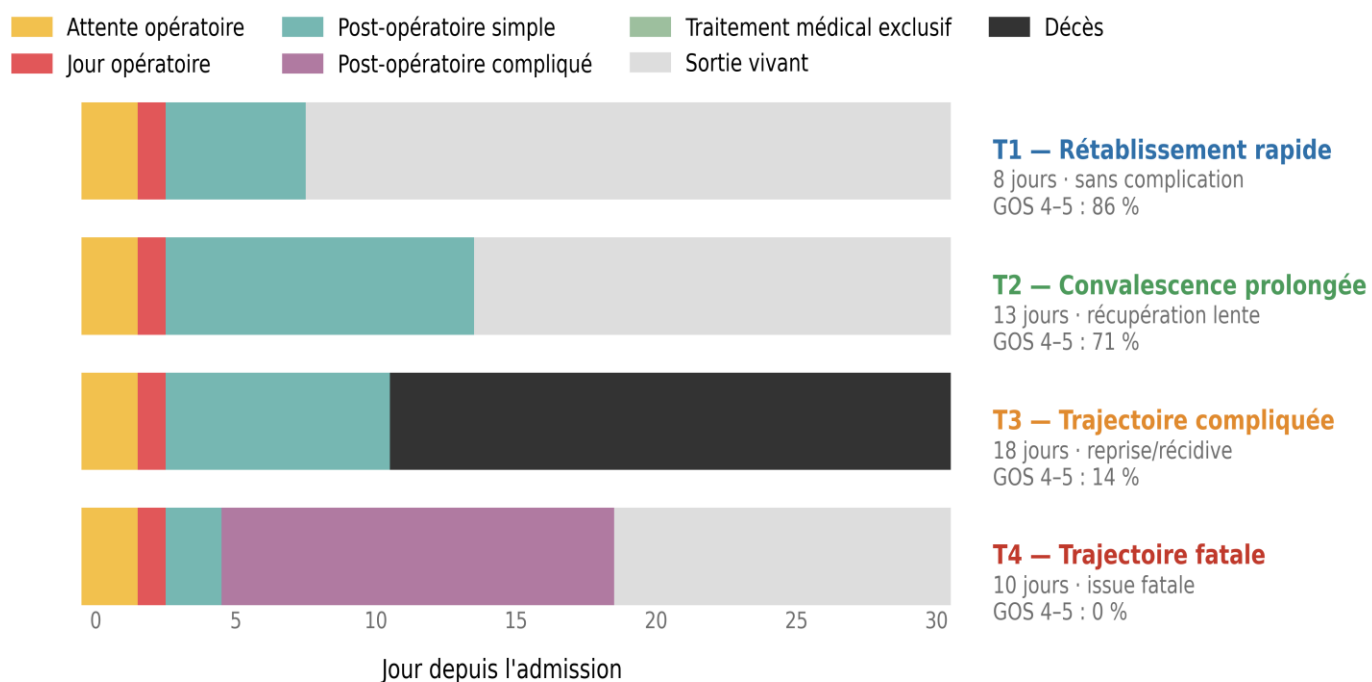


Figure 9. Medoid sequences of the four trajectories and clinical reading. The medoid is the real patient whose sum of dissimilarities to the other members of the group is minimal: it is the most typical representative. The right-hand column recalls the clinical parameters of each group.

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

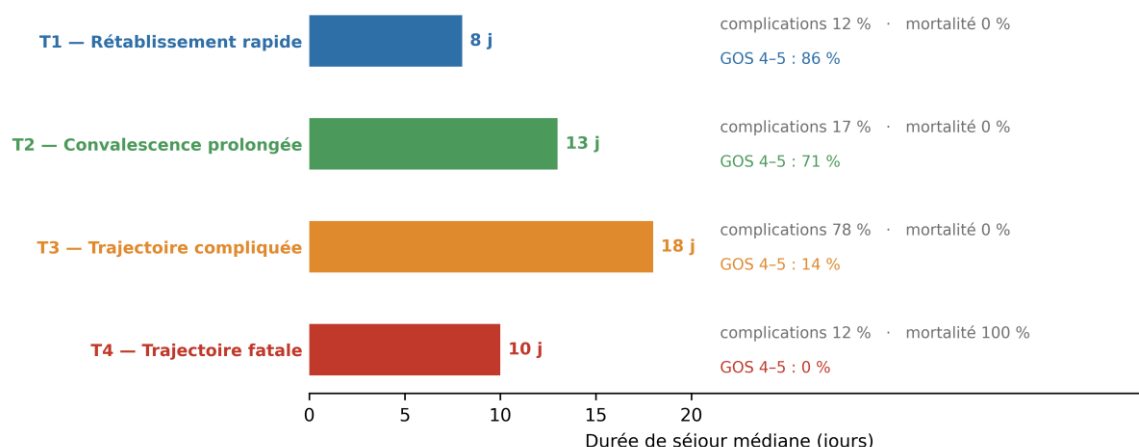


Figure 10. Clinical synthesis of the four trajectories. Median length of stay (bars), proportion of stays with complication, reoperation or recurrence, mortality and proportion of favourable functional outcome, by trajectory. Summarises Table 5.

T1, "rapid recovery", comprises 58 stays (38.7% [31.2-46.7]). The stay is short: median length 8 days (IQR 7-9; 95% CI of the median 7.0-8.5). Admission GCS is 14 (IQR 13-15), no patient has a $GCS \leq 8$, the lesion is the smallest of the cohort (31 mm; IQR 26-36) and the use of surgery the lowest (67.2% [54.4-77.9]). The median patient waits two days, is operated on the second day and is discharged on the ninth. No deaths; GOS 4-5 for 47 of the 55 assessable stays (85.5% [73.8-92.4]).

T2, "prolonged convalescence", comprises 53 stays (35.3% [28.1-43.3]). Admission GCS resembles that of T1 (13; IQR 10-15), but the lesion is larger (38 mm; IQR 31-46), surgery almost systematic (94.3% [84.6-98.1]) and the stay five days longer (13 days; IQR 12-15). This is the group that spends the most time in the simple postoperative state (9.4 days out of 31). No deaths; GOS 4-5 for 35 of the 49 assessable stays (71.4% [57.6-82.2]). What separates T2 from T1: lesion size, the use of surgery (94% versus 67%) and the length of the postoperative stay.

T3, "complicated trajectory", comprises 23 stays (15.3% [10.4-22.0]). These are the most severe patients at admission (median GCS 8, IQR 6-10; 52.2% with a $GCS \leq 8$), with the largest lesions (43 mm; IQR 38-54). They account for 18 of the 36 complications in the cohort (78.3% of the group [58.1-90.3]). The stay is the longest: 18 days (IQR 16-19), i.e. 10 days more than T1 (95% CI of the difference 9.0-11.0). Mean time in the complicated postoperative state reaches 11.6 days out of 31; it follows from the complication and the length of stay, the complicated state being set by convention from the third postoperative day. No deaths, but a heavy functional outcome: median discharge GCS of 10 (IQR 8-12), GOS 4-5 for 3 of the 22 assessable stays (13.6% [4.7-33.3]). This is costly survival.

T4, "fatal trajectory", brings together the 16 deaths (10.7% [6.7-16.6]). Their admission profile is not the most severe: median GCS of 12 (IQR 8-13), 31.3% with $GCS \leq 8$, against 52.2% in T3. On the other hand, the delay before admission is the longest (22 days; IQR 18-30). Fourteen of the sixteen patients were operated on, with a median delay of 2 days: death therefore does not stem from a failure of access to the theatre. Median length of stay: 10 days (IQR 8-13).

To read Table 5, two kinds of variables must be separated. The first are used to build the sequences: length of stay ($\epsilon^2 = 0.746$), admission GCS (0.323) and discharge GCS (0.566), complications ($V = 0.543$), death, use of surgery. They describe the composition of the groups; their differences are expected and do not validate the typology. Functional outcome ($V = 0.656$) depends on them in part, since the GOS comes from sequelae and discharge GCS. Among the variables external to the sequences, only lesion diameter clearly distinguishes the groups ($\epsilon^2 = 0.185$). Symptom-to-admission delay distinguishes them only weakly ($\epsilon^2 = 0.049$; $p = 0.017$). Age ($\epsilon^2 = -0.019$), duration of antibiotic therapy (0.023), operative delay (0.039) and referral status ($V = 0.049$) do not distinguish them.

Delay before admission

Median delay went from 16 days in T1 to 20 in T2, 19 in T3 and 22 in T4 ($p = 0.017$; $\epsilon^2 = 0.049$), with no regular order between groups. The difference in medians between T4 and T1 was 6 days (95% CI 1.5-14.5); we took this pair because it is the most distant, which inflates the difference.

The association is modest and unadjusted. In the Firth regressions, which take GCS and lesion diameter into account, delay is not significantly linked to any group (T4: OR per 10 days 1.70 [0.88-3.21]). It is therefore possible that severity at admission carries most of the link, the GCS being a plausible mediator of delay (a hypothesis examined in a related work on the same cohort). We cannot assert that delay weighs on the shape of the stay independently of severity.

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

What this result does not show. It does not show that a shorter delay would make a patient change trajectory: the CI of the difference in medians is wide and the effect size is small. The causal question was dealt with separately, by g-computation, in an earlier work on this cohort.

Discrepancy analysis

Discrepancy analysis (Table 7) tests the link between a covariate and the set of sequences, without going through a partition. Four covariates were linked to the shape of stays: GCS ≤ 8 (pseudo- $R^2 = 0.184$ [bootstrap 95% CI 0.108-0.299]; pseudo-F = 33.3; $p = 0.001$), radiological mass effect (0.091 [0.053-0.146]; $F = 14.9$; $p = 0.001$), transport by ambulance (0.032 [0.011-0.084]; $F = 4.84$; $p = 0.009$) and symptom-to-admission delay above the median (0.021 [0.002-0.069]; $F = 3.22$; $p = 0.026$). The link with GCS ≤ 8 is expected, the severity channel taking up the admission value. Of ten covariates tested, only GCS ≤ 8 and mass effect pass the Bonferroni threshold (0.005); neither the ambulance ($p = 0.009$) nor the delay ($p = 0.026$) does. Delay was cut at the median of the cohort (19 days), a threshold drawn from the data. The analysis covers the 150 stays, the ambulance being recorded for all.

Table 7. Discrepancy analysis: association between covariates and overall shape of the trajectory

Covariate	Pseudo- R^2 [95% CI]	Pseudo-F	p
Admission GCS ≤ 8	0.184 [0.108–0.299]	33.33	0.001
Radiological mass effect	0.091 [0.053–0.146]	14.88	0.001
Transport by ambulance	0.032 [0.011–0.084]	4.84	0.009
Symptom-to-admission delay > 19 days	0.021 [0.002–0.069]	3.22	0.026
Type of suppuration (brain abscess)	0.013	1.87	0.15
Documented positive culture	0.009	1.36	0.24
Sex	0.009	1.31	0.29
Rural residence	0.006	0.82	0.47
Referred patient	0.005 [0.001–0.033]	0.70	0.55
Age ≥ 40 years	0.001	0.13	0.91

Covariates ranked by decreasing pseudo- R^2 . Pseudo- R^2 : share of the variability of dissimilarities between sequences explained by the covariate (bootstrap 95% CI, 500 resamples). Pseudo-F: analogue of the analysis-of-variance F, computed on the dissimilarities. p: permutation test (1,000 permutations; minimum attainable 0.001). The analysis does not depend on any choice of partition. It covers the 150 stays; documentation of a positive culture (50 stays) and ambulance are recorded for all.

Six covariates were not linked to the shape of stays: referral status (0.005 [0.001-0.033]; $p = 0.55$), age ≥ 40 years (0.001; $p = 0.91$), sex (0.009; $p = 0.29$), documentation of a positive culture (0.009; $p = 0.24$), type of suppuration (0.013; $p = 0.15$) and rural residence (0.006; $p = 0.47$).

Two of these negative results matter, provided we remember that an absence of association is not an absence of effect. The first: being referred by another facility is linked neither to the shape of the stay nor to any of its parameters, although half of the patients are. The interval of the pseudo- R^2 (0.001-0.033) mainly rules out a large effect. The second: documentation of a positive culture (50 stays out of 150, i.e. 43.1% of the samples taken) is not linked to the shape of the sequences ($p = 0.24$), which does not allow us to judge the clinical usefulness of bacteriological information.

The link with the ambulance, significant but weak, is not read as a cause: the ambulance probably signals a severity perceived at referral rather than shaping the stay. We speak of transport by ambulance, without prejudging its level of medicalisation.

Who belongs to which group

One-versus-all Firth regressions (Table 8, Figure 11) refine the picture. A GCS ≤ 8 reduced membership of T1 (OR 0.09; 95% CI 0.00-0.76; $p = 0.022$) as well as of T2 (OR 0.09; 0.02-0.37; $p < 0.001$), and strongly increased that of T3 (OR 8.48; 2.53-31.40; $p < 0.001$) and T4 (OR 5.15; 1.21-22.53; $p = 0.027$). Each additional 10 mm of lesion reduced membership of T1 (OR 0.61; 0.37-0.96; $p = 0.032$). Mass effect increased that of T2 (OR 2.89; 1.12-8.09; $p = 0.028$). For T4, prehospital delay went in the expected direction without being conclusive (OR per 10-day increment 1.70; 0.88-3.21; $p = 0.108$). Referral status was linked to no group (all $p > 0.49$), nor was age.

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

Table 8. Determinants of membership of each trajectory: Firth penalised logistic regressions, one-versus-all

Variable	T1 OR (95% CI)	T2 OR (95% CI)	T3 OR (95% CI)	T4 OR (95% CI)
Symptom-to-admission delay (per 10 days)	0.64 (0.38–1.04)	1.21 (0.76–1.92)	0.98 (0.46–1.94)	1.70 (0.88–3.21)
Admission GCS ≤ 8	0.09 (0.00–0.76)*	0.09 (0.02–0.37)***	8.48 (2.53–31.40)***	5.15 (1.21–22.53)*
Age (per 10 years)	1.02 (0.86–1.20)	1.01 (0.86–1.19)	1.01 (0.80–1.27)	0.96 (0.76–1.20)
Referred patient	1.11 (0.52–2.36)	1.13 (0.54–2.37)	0.71 (0.24–2.00)	0.96 (0.33–2.84)
Lesion diameter (per 10 mm)	0.61 (0.37–0.96)*	1.44 (0.94–2.24)	1.37 (0.75–2.45)	0.91 (0.48–1.64)
Mass effect	0.44 (0.17–1.07)	2.89 (1.12–8.09)*	2.27 (0.45–22.35)	0.51 (0.13–2.14)

Adjusted odds ratios from models including all six variables simultaneously. 95% CI by penalised profile likelihood; p by likelihood ratio test. Firth penalisation corrects the complete separation of GCS ≤ 8 in T1. With 16 events in the smallest group, these estimates are exploratory and must not be used for individual prediction. * $p < 0.05$ · $p < 0.01$ · *** $p < 0.001$.

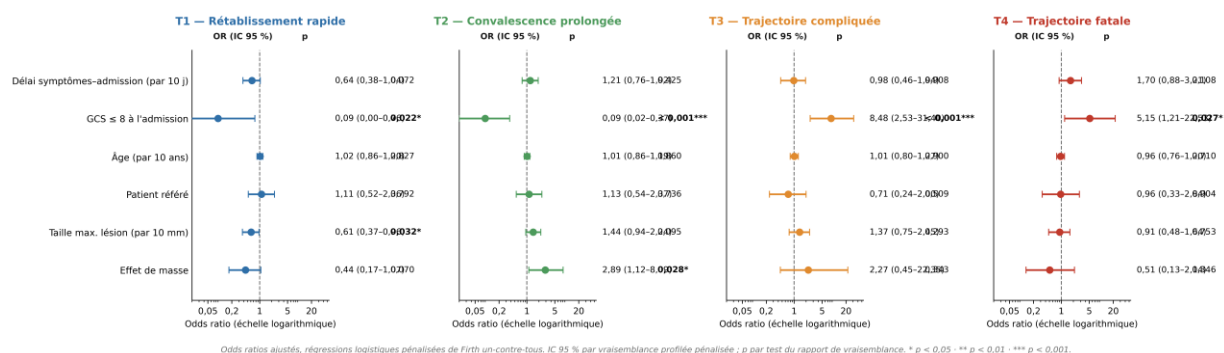


Figure 11. Determinants of membership of each trajectory (one-versus-all Firth regressions). Adjusted odds ratios from models including all six variables simultaneously. 95% CI by penalised profile likelihood; p by likelihood ratio test. Firth penalisation corrects the complete separation of GCS ≤ 8 in T1. With 16 events in the smallest group, these estimates are exploratory and must not be used for individual prediction. * $p < 0.05$ · ** $p < 0.01$ · * $p < 0.001$.**

GCS ≤ 8 is more strongly associated with T3 (OR 8.48) than with T4 (OR 5.15), but the intervals overlap widely, no test compares the two estimates and T4 has only sixteen events for six covariates: this contrast is exploratory. It also stems from the construction of the groups, T4 bringing together all the deaths and T3 none. The risk of death is read more simply: among the 19 stays with GCS ≤ 8 , 5 ended in death (26%), against 11 of 131 (8%) for the others.

Stability of the groups

The bootstrap (Table 9) gives contrasting results. T4 is very stable: mean Jaccard of 0.985 (95% CI 0.938–1.000), above 0.75 in 100% of resamples. T1 reaches 0.741 (0.387–1.000), with 45.7% of resamples above the threshold; T3, 0.672 (0.440–1.000) and 21.0%; T2, 0.599 (0.309–0.983) and 31.3%.

Table 9. Stability of the partition by non-parametric bootstrap (300 resamples)

Trajectory	Mean Jaccard	IC 95 %	Resamples > 0.75	Interpretation
T1 Rapid recovery	0.741	0.387–1.000	45.7 %	Real pattern, blurred boundary with T2
T2 Prolonged convalescence	0.599	0.309–0.983	31.3 %	Continuum rather than group
T3 Complicated trajectory	0.672	0.440–1.000	21.0 %	Real pattern, limiting sample size
T4 Fatal trajectory	0.985	0.938–1.000	100.0 %	Stable group (absorbing state)

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

Hennig procedure: independent clustering of each bootstrap sample, reassignment of all original patients to the bootstrap medoids, then maximal Jaccard index with the reconstituted groups. Thresholds: > 0.75 stable group; 0.60-0.75 real pattern with blurred boundaries; < 0.50 non-reproducible.

T4 is a group in the strong sense, which is not surprising: the absorbing state of death structurally separates it from the others. The typology therefore comprises three survivor profiles and one profile of death. T1 and T3 are real patterns with blurred boundaries. T2 (0.599) falls in the 0.50-0.60 zone that the thresholds leave unqualified: it is a continuum of lengths of stay whose boundary with T1 is arbitrary, to be read as a segment of a continuum rather than as a type. We do not, however, make it an artefact.

Sensitivity analyses

The eleven sensitivity analyses (Table 10, Figure 12) set the scope of the typology. Six give partitions close to the main solution ($ARI \geq 0.70$): horizon reduced to 21 days (AS6, 0.992), exclusion of the ten stays with imputed operative delay (AS7, 0.908), complicated state shifted to the fifth postoperative day (AS2, 0.893), constant substitution costs (AS4, 0.888), Ward hierarchical clustering [50] instead of PAM (AS8, 0.771) and Hamming distance instead of optimal matching (AS10, 0.704). The last two are only moderately robust on the scale of Table 10. Imputation of operative delay and late dating of the complication therefore do not create the structure, without however leaving the partition intact.

The others are less reassuring. Moving the complicated state forward to the second postoperative day (AS1, ARI 0.534) gives a partition that is only moderately robust: the typology is sensitive to this assumption. Three analyses fall below 0.50, which is a limit. The care channel alone (AS3) gives 0.265: the severity channel therefore contributes greatly to the structure. The LCS distance (AS11), applied to both channels with equal weights, looks only at the order of states, not their timing. It also gives 0.265, with the same group sizes as AS3 (79 / 33 / 23 / 15). The order of states being almost fixed (Table 3), the gap with the main solution stems from dates and durations, that is, from what the sequences vary. Finally, average-linkage clustering (AS9) gives 0.337 with a degenerate partition (one group of 119 stays and three small groups of 15, 13 and 3): this is a weakness of the algorithm on this type of data, not an instability of the structure.

Aligning the sequences on symptom onset rather than on admission (AS5) gives an ARI of 0.370 and a partition dominated by the length of the prehospital delay. This is not a contradiction: a different object is then described, the pathway from the onset of illness and not from entry to hospital.

Table 10. Sensitivity analyses: influence of assumptions on the partition, ranked from most robust to most sensitive

Sensitivity analysis	ARI	ASW	Group sizes	Reading
AS6 Horizon reduced to 21 days	0.992	0.372	58 / 53 / 24 / 15	Negligible influence
AS7 Exclusion of imputed operative delays (n = 140)	0.908	0.365	53 / 51 / 22 / 14	Imputation does not create the structure
AS2 Complicated state from postoperative D+5	0.893	0.381	56 / 55 / 23 / 16	Low influence
AS4 Constant substitution costs	0.888	0.366	58 / 53 / 23 / 16	Little dependent on the choice of costs
AS8 Ward hierarchical clustering	0.771	0.405	60 / 59 / 16 / 15	Close partition, with no 23-stay group equivalent to T3
AS10 Hamming distance (combined channels)	0.704	0.371	59 / 44 / 31 / 16	Structure partly recovered without optimal matching
AS1 Complicated state from postoperative D+2	0.534	0.427	74 / 33 / 27 / 16	Real limitation: influential assumption
AS5 Alignment on symptom onset	0.370	0.396	66 / 43 / 25 / 16	Describes a different object; partition dominated by delay
AS9 Average-linkage hierarchical clustering	0.337	0.460	119 / 15 / 13 / 3	Degenerate partition: weakness of the algorithm

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

Sensitivity analysis	ARI	ASW	Group sizes	Reading
AS3 Care channel alone (without severity channel)	0.265	0.587	79 / 33 / 23 / 15	The severity channel structures the partition
AS11 LCS distance (two channels; order without timing)	0.265	0.582	79 / 33 / 23 / 15	Same ARI and group sizes as AS3; the order being almost fixed, the gap stems from dates and durations

ARI: adjusted Rand index between the sensitivity partition and the main partition (1.000 = identical; 0.000 = chance agreement). Each analysis redoes the whole chain. Reading: ARI ≥ 0.75 robust; 0.50-0.75 moderately robust; < 0.50 sensitive. Analyses are numbered AS1 to AS11 so as not to confuse them with supplements S1 to S8.

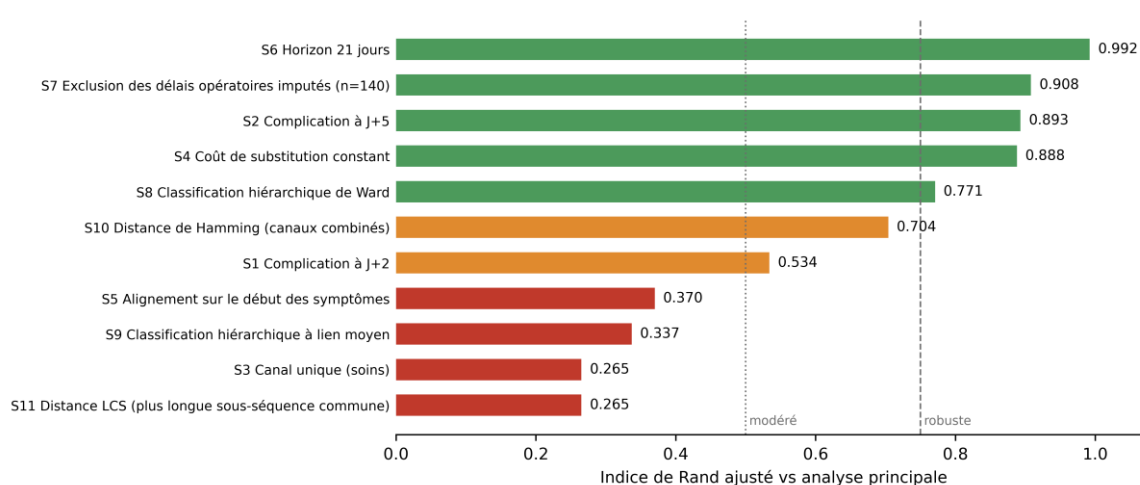


Figure 12. Robustness map of the eleven sensitivity analyses. ARI between each sensitivity partition and the main partition, in decreasing order. Colours: robust (≥ 0.75), moderately robust (0.50-0.75), sensitive (< 0.50). The two vertical lines mark these thresholds.

Table 11. Summary of internal validation

Level assessed	Done	Method	Result
Robustness of the dissimilarity metric	Yes	Constant costs · Hamming · LCS	Robust to costs (AS4, 0.888); moderately robust to Hamming (AS10, 0.704); clear gap with LCS (AS11, 0.265)
Robustness of the clustering algorithm	Yes	Ward · average linkage	Robust to Ward (AS8, 0.771); degenerate partition with average linkage (AS9)
Robustness of the groups	Yes	Hennig bootstrap, 300 resamples	1 stable group, 2 real patterns, 1 continuum
Robustness of the associations	Yes	Bootstrap CIs of pseudo- R^2 · profile likelihood	Main associations confirmed; negative results confirmed
Robustness to reconstruction assumptions	Partial	AS1, AS2, AS6, AS7; GCS pivot date not tested	Horizon and imputation: low influence; dating of complications: sensitive at D+2
External validation	No	—	Main limitation; multicentre programme proposed

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

Level assessed	Done	Method	Result
Reproducibility	Yes	Code, dictionary, reconstruction rules	Supplements S1-S8
Reporting compliance	Yes	STROBE checklist	Supplement S5

Validity was assessed separately at four levels of the analytical chain. No external validation was carried out: this is the first limitation of this work.

DISCUSSION

In brief

Stays for ISI are varied but fall into few profiles, of which only one, death, is truly stable. They remain dispersed from about the second to the twelfth day, which mainly reflects the dispersion of lengths of stay. Delay before admission has only a modest, unadjusted link with the shape of the stay; referral status has none. Table 12 separates what is observed, interpreted or simply assumed.

Table 12. Synthesis: questions, answers and level of evidence

Question	Answer	Status
How many forms of pathway?	Four, explaining 77.8% of the variability of the dissimilarities	Observation
When do pathways diverge?	High entropy from D2 to D12; maximum of 0.717 on D10 (D2: 0.701)	Observation (reflects lengths of stay)
Why at this time?	Dispersion of discharge dates and D+3 convention; clinical window not testable	Interpretation
Main clinical determinant	GCS ≤ 8 (pseudo- $R^2 = 0.184$, partly constitutive)	Observation
Organisational determinant	Symptom-to-admission delay: modest, unadjusted association, absent from adjusted models (pseudo- $R^2 = 0.021$)	Exploratory observation
Does coma predict death?	Death: 26% if GCS ≤ 8 versus 8%; ORs of 8.48 (T3) and 5.15 (T4) indistinguishable	Observation
Main negative result	Referral status is associated with no form of pathway	Observation
Clinical message	Codify postoperative monitoring (does not follow from the data)	Hypothesis to evaluate
Organisational message	The distribution between trajectories as a service indicator	Hypothesis to evaluate
Transferability	Method transferable; typology to be validated elsewhere	Hypothesis to evaluate

This table summarises the messages of the manuscript and their level of evidence. "Observation": directly measured result; "interpretation": reading compatible with the data but not demonstrated; "hypothesis": proposal to be tested.

The four groups

We retained four groups because they read well clinically, the silhouette being lowest at $k = 4$ (Table 4). One can see in them the crossing of two dimensions: neurological severity at admission and the occurrence of a complication. A mildly ill patient without complication follows T1 or T2, the difference lying in lesion size, the use of surgery and the length of convalescence. A severely ill patient with a complication follows T3. T4 groups the deaths. As the sequences are functions of length of stay, complication and outcome, this reading follows in part from the construction; we do not propose it for other neurosurgical emergencies.

It remains a hypothesis. The low stability of T2 suggests, moreover, that the boundary between the two forms without complication is gradual: a continuous speed of recovery rather than a dichotomy.

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

Dispersion around the tenth day

This maximum is read first in the light of the construction. The complicated state is set by convention from the third postoperative day, which mechanically raises entropy on the following days: shifting it to the fifth day leaves the maximum unchanged, but moving it forward to the second day clearly changes the partition (ARI 0.534). Discharge occurs at a median of the eleventh day; at that date, discharged patients sit alongside patients still in hospital, so that entropy reflects the dispersion of lengths of stay as much as that of clinical states. Finally, the maximum (0.717 on day 10) exceeds the first peak (0.701 on day 2) by only 0.016.

One can imagine that the end of the first week and the second postoperative week are when re-accumulated collections, secondary infectious complications and late neurological deteriorations appear, and that favourable and unfavourable courses then separate. As the registry does not date complications, we cannot test this idea. For a patient operated on the second day, the tenth day of hospitalisation is the eighth postoperative day.

These data do not justify shifting the improvement effort from upstream to postoperative monitoring. The usual targets (consultation delay, referral chain, triage at reception) remain relevant; postoperative monitoring is a lead to test, not a result.

Initial severity, complicated survival and death

A low GCS is associated both with death and with survival with heavy sequelae. Of the 19 stays with $GCS \leq 8$, 5 ended in death (26%, against 8% for the others), 12 followed the complicated trajectory and 2 the prolonged convalescence; none joined rapid recovery. The deaths being grouped in T4 by construction, the difference between the ORs of T3 and T4 does not mean that the GCS predicts death less well: it describes how severely ill patients are distributed between two profiles, one of which excludes death. T3 brings together the lowest GCS and the largest lesions, had no deaths, but a very impaired functional outcome (GOS 4-5: 3 stays out of 22).

This agrees with the prognostic model built on this cohort: moderate discrimination (AUC 0.70, i.e. 0.66 after optimism correction) and $GCS < 9$ the only factor independently associated with death (aOR 3.79; 95% CI 1.11-12.92). To inform families, a $GCS \leq 8$ at admission announces both an increased risk of death (26% versus 8%) and, in survivors, an outcome that is most often heavy: in the main article of the series, none of the 17 patients with $GCS < 9$ whose GOS was assessable reached a GOS of 4-5, and 12 of the 19 stays with $GCS \leq 8$ followed T3, the trajectory that consumes the most hospital days. The CIs of the two ORs overlap, and these observations must not be used to predict the fate of an individual patient.

Delay before admission

On this cohort, a link between delay and mortality has already been reported, with a g-computation estimate of avoidable mortality under different shortening scenarios. Here, delay, cut at the median, explains only a small share of the variability of the sequences (2.1%; unadjusted $p = 0.026$), and the link disappears when adjusting for GCS and lesion size. We therefore cannot conclude that delay marks the shape of the stay independently of severity at admission. The pseudo- R^2 should have been adjusted for GCS; we did not do so.

Referral status explains nothing

The absence of a link between referral status and the shape of the stay contradicts a common idea, including in our earlier work on this cohort, where the analysis of system performance designated referral as a limiting domain. In the main article of the series, a symptom-to-admission delay above 19 days was associated with referral status (aOR 1.96; 95% CI 1.00-3.89) and with rural residence (aOR 2.16; 95% CI 1.12-4.24). Referral therefore lengthens the wait before admission, without this showing in the shape of the hospital stay.

Several explanations are possible. The referral chain may act too late to matter: when the median delay reaches 19 days, a few days of circulation between facilities no longer change the shape of the stay, and the delay would build up upstream, at the time of first contact with care [51]. The referral variable, being binary, distinguishes neither the number of facilities passed through nor the time spent in each. Finally, power may be insufficient: the CI of the pseudo- R^2 (0.001-0.033) rules out a large effect, not a weak one.

These explanations are not mutually exclusive and call for the same response: to document the prehospital phase as finely as the hospital phase. A sequence starting at first contact with care would settle the matter. Our analysis aligned on symptom onset gives a first approximation, with a partition dominated by the length of wandering.

The blind spot: intensive care

Our conceptual scheme provided for an intensive care state between surgery and course. We could not reconstruct it: the registry records neither admission to intensive care, nor intubation, nor ventilation, nor vasopressors, and systematic examination of the free-text fields found no mention of them. The severity channel is only an imperfect substitute: it captures the patient's state, not the intensity of care.

We report this gap rather than filling it with a hazardous reconstruction, and we regard it as a result. In a department that handles about fifteen ISI a year, one patient in eight arriving with a $GCS \leq 8$, intensive care leaves no written trace. It is therefore impossible

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

to assess whether critical care needs are met, although this is one of the known bottlenecks of neurosurgery in sub-Saharan Africa. Our first recommendation is thus documentary before being clinical.

What the literature says

NORTH AMERICA. Vanasse and colleagues, in Quebec, proposed a multidimensional model of care pathways [52] and applied it to the course after a first hospitalisation for COPD [15], to the pathways preceding a first diagnosis of schizophrenia [16] and to antipsychotic treatment trajectories [18]. Other Canadian teams studied inequalities in pathways after a diagnosis of diabetes [17] and care trajectories in arthritic conditions [53].

EUROPE. Sequence analysis has described prenatal pathways [20] and multiple sclerosis pathways in France [54], as well as a chronological mapping of 308 conditions in the English health system [55]. A large-scale French analysis grouped pathways after an intensive care stay according to the number of days spent at home [56].

Neighbouring approaches. Process mining pursues a similar aim from time-stamped event logs [46, 47]: characterisation of hospital pathways [57-59], comparison of subpopulations, monitoring of equity of access [60], temporal phenotyping [57, 61]. It assumes a hospital information system, which our context does not offer, but its questions largely overlap with ours [62, 63]. A scoping review listed the available methods for identifying sequences of use in health data [29], and a systematic review examined the methodological practices of sequence analysis applied to clinical pathways [30].

Africa and neurosurgery. We found no study applying sequence analysis to intracranial suppurative infections or to a neurosurgical department in sub-Saharan Africa; this search must be updated before submission. For other neurological conditions, notably head injury, sequence analyses exist, based on French hospital and health insurance data [64]. African work on ISI remains, for the most part, descriptive [3-6]. This methodological lag is specific neither to this subject nor to this continent, but it is particularly marked there.

Clinically, our cohort resembles the published African series [3-6]: predominant brain abscess, frequent ENT and stomatological portals of entry, low-yield cultures, long prehospital delays, mortality between 10 and 25% [2, 4, 65]. Our 10.7% lies at the low end of this range, with a reservation: the cohort counts only the patients who reached the department, and deaths occurring before are not counted. Antibiotic therapy (seven weeks at the median) and surgical drainage in more than eight cases out of ten correspond to the recommended regimens [66-69].

FOR PRACTICE

What follows remains to be evaluated.

Identify at reception the patients at risk of a complicated trajectory. A $GCS \leq 8$ is the clearest marker (12 of the 23 stays of T3). It makes it possible to anticipate a long stay and a probable complication, but not to predict the trajectory of a given patient.

Code postoperative monitoring. A codified daily neurological examination, control imaging in case of stagnation and explicit criteria for reoperation are reasonable in all cases. They do not follow from our data, which do not date complications; their effect on the distribution of trajectories remains to be shown, and their cost was not assessed.

Talk to families without minimising the initial coma. A $GCS \leq 8$ announces an increased risk of death and, in survivors, an outcome that is most often heavy (see above). The trajectories are population profiles intended to better organise care, not individual diagnostic categories.

Assessing the quality of the service

The classic indicators of a neurosurgical department are mortality, complication rate and mean length of stay [70, 71]. In our setting, mortality concerns only one stay in ten, and the other two summarise the stay in a single figure. The distribution of stays between trajectories could be a more sensitive composite indicator: a shift from T3 to T1 or T2, at comparable admission severity, would reflect progress that a mortality stable at 10% would not see. The logic resembles the care cascades described for HIV [72, 73], which measure a pathway rather than a single result.

Indicators drawn from this analysis, such as the proportion of stays in the complicated trajectory or the time spent in the complicated state, could be added to them. They depend on the conventional setting of the complicated state and on the moderate stability of the typology: we propose them only as a lead, in the continuation of the quality indicator framework proposed on this cohort.

This logic (collect, analyse, improve, collect again) is that of learning health systems [74, 75]. It requires a registry kept consistently: the quality of the reconstructed trajectories depends on the completeness of the longitudinal data. A registry is here a methodological investment as much as a documentary one [76].

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

Beds and hospital days

Without an economic analysis, an order of magnitude suffices. T3 represents 15.3% of stays but, from the mean times in Table 6, about 23% of hospital days (about 19 days per stay on average, against 9 for T1). In a department with limited capacity, reducing the share of T3 a little would free beds. A medico-economic evaluation would be justified; we did not carry it out.

Relevance for global neurosurgery

Work on neurosurgery in sub-Saharan Africa has documented the lack of operative capacity, neurosurgeons and technical facilities [77-82], in the wake of the Lancet Commission on Global Surgery [83] and of work on global neurosurgical capacity [84-86]. It has more rarely looked at what happens inside the departments that exist.

In our setting, cross-sectional imaging was available for all included stays (a condition of inclusion, which says nothing about the delay of access), median operative delay was two days and seven operated patients in ten had been operated on within 48 hours. Lacking a referenced comparator, we do not conclude that these indicators equal those of better-resourced settings: for suppurative collections, two days is not necessarily satisfactory. The nineteen-day delay before admission and the quarter of stays marked by a complication, reoperation or recurrence are the two other striking findings.

These observations are of interest for the evaluation of neurosurgical systems. Capacity indicators (number of neurosurgeons, of scanners, operative volume [87, 88]) measure what a system can do, not the actual pathway of patients. National surgical, obstetric and anaesthesia plans, which guide surgical planning in several African countries [89-91], would benefit from adding pathway measures. The literature on the quality of health systems also distinguishes access to care from the quality of the care process [71, 92, 93].

These principles could serve other resource-limited neurosurgical departments, to compare not only results but the structure of pathways. For this, the alphabet of states would first have to be harmonised and validated locally.

When to use the method

It is suitable when the question concerns the organisation of the stay over time, when several states succeed one another and when dated landmarks exist for at least three of them.

It is not suitable for predicting a single event (a prognostic model is then more appropriate), nor when all transitions are observed and dated (a multistate model is more economical in assumptions), nor when the states rest on more than two or three untestable assumptions. This is the case here (five assumptions, of which the GCS pivot date was not tested), hence the caution of our conclusions.

The minimum seems to us to be: declare the alphabet of states in advance, publish the reconstruction rules, test each assumption by a sensitivity analysis and report the stability of the groups. Figure 13 contrasts the two analytical frameworks; an eight-step guide drawn from this work is provided in S7.

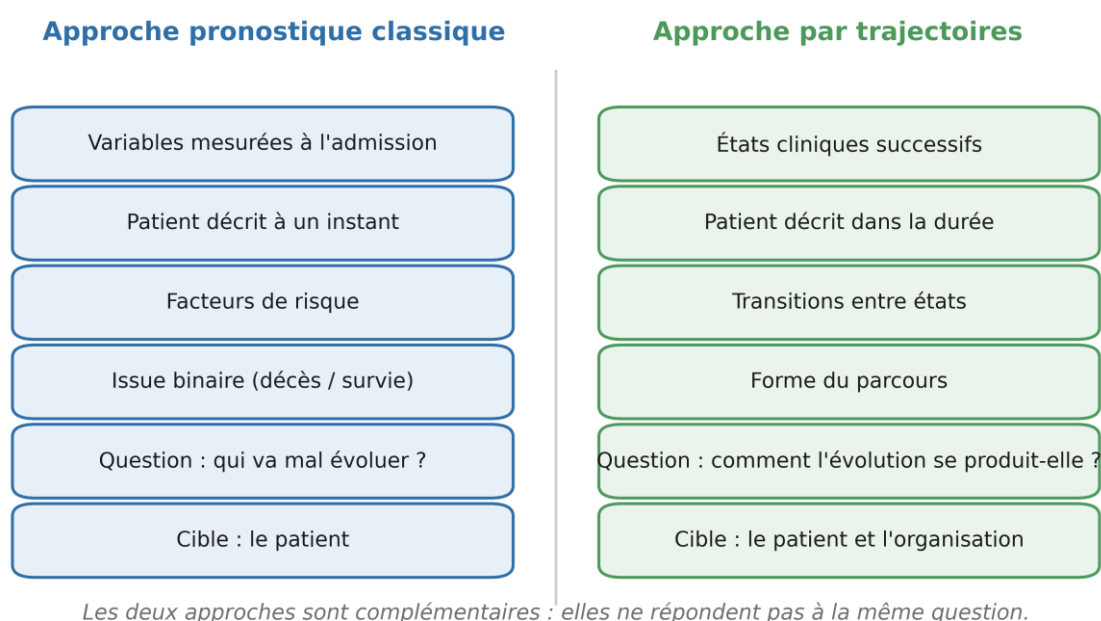


Figure 13. From variable-based analysis to trajectory-based analysis. The two analytical frameworks. Left, the conventional prognostic approach: variables measured at one time point, binary outcome. Right, the trajectory approach: succession of states, shape of the stay. The two answer different questions and complement each other.

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

Strengths

This study appears to be the first application of sequence analysis to intracranial suppurative infections; we cannot exclude non-indexed work. The registry is detailed (112 variables over ten years). The alphabet of states and the reconstruction rules were fixed before clustering. Eleven sensitivity analyses and a bootstrap of 300 resamples inform the stability of the typology, and Firth penalisation suits small numbers and the complete separation observed.

LIMITATIONS

No external validation. The typology comes from a single-centre cohort of 150 stays and has not been reproduced anywhere else: these are the trajectories of this department, with its organisation, resources and practices.

Moderate stability. Three groups out of four do not reach the bootstrap reproducibility threshold, and T2 is a continuum rather than a group. The typology is an exploratory descriptive framework, not an established classification.

Statistical groups. Trajectories are not necessarily natural entities: they are groupings that depend on the metric, the alphabet of states and the algorithm. The sensitivity analyses show to what extent.

Reconstruction assumptions. The five assumptions (Table 1) weigh unequally: imputation of operative delay and the horizon have little influence; the conventional dating of complications clearly changes the partition in its earliest version; the GCS pivot date was not tested.

Constitutive variables. Length of stay, complication, GCS and death build the sequences and cannot validate them. The GOS, drawn from sequelae, is not an independent measure. The sequences carry only seven pieces of information from the registry, and the order of states is imposed: the analysis describes dates and durations more than a chain of events.

Structure of the partition. The silhouette of the retained solution (0.378) indicates a weak structure, and the choice of four groups rests on readability. Four sensitivity analyses out of eleven give an ARI < 0.50.

Adjustment and dependence. The pseudo- R^2 of delay was not adjusted for GCS. Two patients contributed two stays, not independent in the permutation tests; the expected effect is negligible.

Intensive care. It is absent from the registry (see above).

Partly retrospective collection. It exposes to information bias, for example a complication under-recorded in patients whose course was quickly favourable.

Small numbers. With sixteen stays in T4 and twenty-three in T3, the Firth models are exploratory and must not be used to predict.

No correction for multiple comparisons. The p values are to be read as descriptive. The analysis protocol, time-stamped before the analyses, was not publicly pre-registered.

Generalisability

What should transpose: the reconstruction method, the alphabet of states, the validation approach, and the idea that an ordinary clinical registry is enough to produce a sequence analysis.

What will probably not transpose: the proportion of the four trajectories, the position of the entropy maximum and the extent of the prehospital delay depend on context. With a dedicated neurosurgical intensive care unit, a critical care channel would change the structure of the sequences and probably the typology. If the prehospital delay were three days rather than nineteen, its link with the shape of the stay might be different. With permanently accessible MRI, complications would be detected earlier and the dispersion window might shift.

What this study does not show

It does not show that a trajectory is written from admission, nor that a shorter delay would make a patient change trajectory, nor that delay weighs on the shape of the stay independently of severity. It does not say that acting on the second week would improve outcome, nor that the dispersion of stays peaks at a precise moment beyond what the distribution of lengths implies. It does not prove that these four trajectories exist elsewhere, nor that initial severity truly distinguishes complicated survival from death, the CIs of the two estimates overlapping.

It shows that, in this cohort, stays reconstructed according to the declared rules fall into four forms, including a profile of death; that these forms differ by the variables that build them and by lesion size; and that the overall shape of the sequences is linked to GCS ≤ 8 and to mass effect.

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

PERSPECTIVES

Step 1, document. Add to the registry the date of admission to intensive care and its duration, the date of complications and the number of facilities passed through before admission. These four variables would lift most of the reconstruction assumptions.

Step 2, go upstream. Extend the sequence back to first contact with care, to distinguish delay in seeking care from referral delay.

Step 3, validate elsewhere. Repeat the analysis in several Ivorian then West African centres, with a common alphabet of states. The four trajectories will have to be found there before serving as a reference.

Step 4, compare. Compare the structure of pathways with that of better-resourced centres, on harmonised states.

Step 5, predict. Test whether the group can be predicted from admission, with multistate models or supervised learning. This is a lead, not a result.

Step 6, intervene. Evaluate, by a hybrid effectiveness-implementation design, the effect of codified postoperative monitoring on the distribution of stays between trajectories.

CONCLUSIONS

In this cohort of 150 stays for intracranial suppurative infection, day-by-day reconstruction of the pathway distinguishes four profiles: rapid recovery for nearly four patients in ten, prolonged convalescence for a third, a complicated course with costly survival for one in six to seven, and in-hospital death for one in ten.

These profiles are distinguished mainly by what builds them (length of stay, complication, outcome) and by lesion size. The dispersion of stays is maximal from the second to the twelfth day, which largely reflects the dispersion of lengths of stay. Delay before admission has only a modest, unadjusted link with the shape of the stay; referral status has none.

In terms of method, a sequence analysis is feasible with an ordinary clinical registry, without a hospital information system, provided the reconstruction rules are written down and their influence tested. The limits are clear: three groups out of four are below the reproducibility threshold, the order of events is imposed for lack of complication dates, and intensive care, although central for a comatose patient, remains undocumentable.

Trajectory analysis does not replace prognostic approaches: it answers a different question. A prognostic model estimates a risk; sequence analysis describes how durations, complications and outcomes are distributed. Multicentre validation and a registry that dates complications and intensive care are needed before making it a framework for complex neurosurgical pathways.

DECLARATIONS

Ethics approval. Bouaké University Hospital did not have a constituted ethics committee during the study period; the study was conducted after authorisation by the head of the neurosurgery department and the hospital management, in accordance with the principles of the Declaration of Helsinki. It has since received approval from the Institutional Ethics Committee of Bouaké University Hospital (approval no. CEI/CHU-BKE/136/2026).

Consent. Informed consent was obtained for prospective inclusions (2025-2026); a waiver of consent was granted for the retrospective phase, the data having been anonymised.

Data and code availability. The dictionary of the state alphabet, the reconstruction rules, the cost matrices and the analysis code are given in the supplementary material (S2 to S4, S6 to S8). Anonymised individual data are available from the corresponding author on reasoned request, subject to the agreement of the ethics committee.

Author contributions. All authors made a substantial contribution to this work, whether in the conception of the study, data collection, analysis and interpretation of the results, or the drafting and critical revision of the manuscript. All have read and approved the submitted version and take collective responsibility for it.

Overlap with earlier work. This analysis concerns the same cohort as several works of the series: epidemiological profile and prognostic factors, bedside prognostic score, prehospital delay cascade, quality indicators, unsupervised phenotypes, neurosurgical system performance index, triage score, counterfactual analysis of delay. Its question, its unit of analysis (the daily sequence rather than the patient), its method and its results are distinct and have not been published elsewhere. The detailed declaration of related publications, with their status, is enclosed in the submission file.

Transparency statement. The dictionary of states and the reconstruction rules were fixed before classification (S2). The analysis protocol (S1) was finalised and time-stamped before the construction of the sequences, the computation of distances and the classification; it is provided as a supplement, without public pre-registration. Each modelling assumption was subjected to a sensitivity analysis, except the GCS pivot date and the two imputations of missing data. Results unfavourable to the typology are reported as fully as the others.

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

Conflict of interest

The authors declare no conflict of interest.

Funding statement

No specific funding.

Abbreviations

ARI: adjusted Rand index · ASW: mean silhouette width · CI: confidence interval · COPD: chronic obstructive pulmonary disease · CT: computed tomography · GCS: Glasgow Coma Scale · GOS: Glasgow Outcome Scale · IQR: interquartile range · ISI: intracranial suppurative infection · LCS: longest common subsequence · MRI: magnetic resonance imaging · OM: optimal matching · OR: odds ratio · PAM: partitioning around medoids · V: Cramér's V · ε^2 : epsilon squared

Supplementary material

S1 — Analysis protocol: research question, unit of analysis, alphabet of states fixed before classification, rules for choosing the number of groups, planned sensitivity analyses. Finalised and time-stamped before the construction of the sequences; without public pre-registration.

S2 — Dictionary of states and reconstruction rules: definition of the seven care states and the three severity bands, daily assignment rules, handling of absorbing states, five modelling assumptions, annotated example of a sequence.

S3 — Code and environment: construction of sequences, substitution costs, multichannel optimal matching, PAM, discrepancy analysis with permutations, Firth regression, Hennig bootstrap; versioned environment file and reproduction instructions.

S4 — Cost matrices: substitution matrix of the care channel (derived from transition rates), ordinal matrix of the severity channel, insertion-deletion cost and justification.

S5 — Annotated STROBE checklist.

S6 — Methodological notes: formalisation of multichannel optimal matching, time horizon, comparison of metrics, choice of PAM, stability thresholds, limits of absorbing states, frequent errors in sequence analysis on clinical data.

S7 — Eight-step guide to conducting a sequence analysis on a neurosurgical cohort: define the unit of analysis, build the alphabet of states, set the horizon, reconstruct the sequences, compute dissimilarities, choose the algorithm, validate stability, interpret the trajectories.

S8 — Statistical appendix: optimal matching distance, full substitution matrix, medoid pseudo-code, Hennig bootstrap algorithm, discrepancy pseudo- R^2 , Firth penalised likelihood, formulas for entropy, the Jaccard index and the adjusted Rand index.

Table S1 — Complete descriptive characteristics of the cohort (112 variables, 12 sections).

Table S2 — The ten most frequent state sequences and their counts.

Table S3 — Distribution of the Glasgow Outcome Scale and sequelae by trajectory.

Table S4 — Cross-concordance between the main partition and each of the eleven sensitivity partitions.

Table S5 — Characteristics of the ten stays with imputed operative delay.

REFERENCES

- 1) Brouwer MC, Tunkel AR, McKhann GM, van de Beek D. Brain abscess. *N Engl J Med*. 2014;371:447-56.
- 2) [Brouwer MC, Coutinho JM, van de Beek D. Clinical characteristics and outcome of brain abscess: systematic review and meta-analysis. *Neurology*. 2014;82:806-13.
- 3) Nathoo N, Nadvi SS, van Dellen JR, Gouws E. Intracranial subdural empyemas in the era of computed tomography: a review of 699 cases. *Neurosurgery*. 1999;44:529-35.
- 4) Nathoo N, Nadvi SS, Narotam PK, van Dellen JR. Brain abscess: management and outcome analysis of a computed tomography era experience with 973 patients. *World Neurosurg*. 2011;75:716-26.
- 5) Kanu OO, Nnoli C, Olowoyeye O, Arigbabu S. Infratentorial brain abscess: management challenges in a resource-limited setting. *Niger J Clin Pract*. 2012;15:228-32.
- 6) Sanoussi S, Rabiou MS, Bawa KM, Kelani A, Sani R, Chaibou MS. Les suppurations intracrâniennes : aspects épidémiologiques, diagnostiques et thérapeutiques à Niamey. *Afr J Neurol Sci*. 2011;30.
- 7) Studer M, Ritschard G. What matters in differences between life trajectories: a comparative review of sequence dissimilarity measures. *J R Stat Soc Ser A Stat Soc*. 2016;179:481-511.
- 8) Abbott A, Hrycak A. Measuring resemblance in sequence data: an optimal matching analysis of musicians' careers. *Am J Sociol*. 1990;96:144-85.
- 9) Abbott A, Tsay A. Sequence analysis and optimal matching methods in sociology: review and prospect. *Sociol Methods Res*. 2000;29:3-33.

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

- 10) Aisenbrey S, Fasang AE. New life for old ideas: the second wave of sequence analysis bringing the course back into the life course. *Sociol Methods Res.* 2010;38:420-62.
- 11) Gabardin A, Ritschard G, Müller NS, Studer M. Analyzing and visualizing state sequences in R with TraMineR. *J Stat Softw.* 2011;40:1-37.
- 12) Liao TF, Bolano D, Brzinsky-Fay C, Cornwell B, Fasang AE, Helske S, et al. Sequence analysis: its past, present, and future. *Soc Sci Res.* 2022;107:102772.
- 13) Ritschard G, Studer M, editors. *Sequence analysis and related approaches: innovative methods and applications.* Cham: Springer; 2018.
- 14) Cornwell B. *Social sequence analysis: methods and applications.* Cambridge: Cambridge University Press; 2015.
- 15) Vanasse A, Courteau J, Courteau M, Benigeri M, Chiu YM, Dufour I, et al. Healthcare utilization after a first hospitalization for COPD: a new approach of state sequence analysis based on the 6W multidimensional model of care trajectories. *BMC Health Serv Res.* 2020;20:177.
- 16) Vanasse A, Courteau J, Courteau M, Roy MA, Stip E, Fleury MJ, et al. Multidimensional analysis of adult patients' care trajectories before a first diagnosis of schizophrenia. *Schizophrenia (Heidelb).* 2022;8:52.
- 17) McKay R, Letarte L, Lebel A, Quesnel-Vallée A; TORSADÉ Cohort Working Group. Exploring social inequalities in healthcare trajectories following diagnosis of diabetes: a state sequence analysis of linked survey and administrative data. *BMC Health Serv Res.* 2022;22:131.
- 18) Brodeur S, Vanasse A, Courteau J, Courteau M, Stip E, Fleury MJ, et al. Antipsychotic utilization trajectories three years after initiating or reinitiating treatment of schizophrenia: a state sequence analysis approach. *Acta Psychiatr Scand.* 2022;145:469-80.
- 19) Vanoli J, Nava CR, Airoldi C, Ucciero A, Salvi V, Barone-Adesi F. Use of state sequence analysis in pharmacoepidemiology: a tutorial. *Int J Environ Res Public Health.* 2021;18:13398.
- 20) Le Meur N, Gao F, Bayat S. Mining care trajectories using health administrative information systems: the use of state sequence analysis to assess disparities in prenatal care consumption. *BMC Health Serv Res.* 2015;15:200.
- 21) Nguena Nguefack HL, Pagé MG, Katz J, Choinière M, Vanasse A, Dorais M, et al. Trajectory modelling techniques useful to epidemiological research: a comparative narrative review of approaches. *Clin Epidemiol.* 2020;12:1205-22.
- 22) Nagin DS, Odgers CL. Group-based trajectory modeling in clinical research. *Annu Rev Clin Psychol.* 2010;6:109-38.
- 23) Piccarreta R, Studer M. Holistic analysis of the life course: methodological challenges and new perspectives. *Adv Life Course Res.* 2019;41:100251.
- 24) Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health.* 2011;38:65-76.
- 25) Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci.* 2009;4:50.
- 26) Curran GM, Bauer M, Mittman B, Pyne JM, Stedler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. *Med Care.* 2012;50:217-26.
- 27) Plsek PE, Greenhalgh T. Complexity science: the challenge of complexity in health care. *BMJ.* 2001;323:625-8.
- 28) Braithwaite J, Churrua K, Long JC, Ellis LA, Herkes J. When complexity science meets implementation science: a theoretical and empirical analysis of systems change. *BMC Med.* 2018;16:63.
- 29) Flothow A, Novelli A, Sundmacher L. Analytical methods for identifying sequences of utilization in health data: a scoping review. *BMC Med Res Methodol.* 2023;23:212. <https://doi.org/10.1186/s12874-023-02019-y>
- 30) Mathew S, Peat G, Parry E, Singh Sokhal B, Yu D. Applying sequence analysis to uncover 'Real-World' clinical pathways from routinely collected data: a systematic review. *J Clin Epidemiol.* 2024;166:111226. <https://doi.org/10.1016/j.jclinepi.2023.111226>
- 31) von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet.* 2007;370:1453-7.
- 32) Lesnard L. Setting cost in optimal matching to uncover contemporaneous socio-temporal patterns. *Sociol Methods Res.* 2010;38:389-419.
- 33) Elzinga CH, Studer M. Spell sequences, state proximities, and distance metrics. *Sociol Methods Res.* 2015;44:3-47.
- 34) Hollister M. Is optimal matching suboptimal? *Sociol Methods Res.* 2009;38:235-64.
- 35) Wu LL. Some comments on sequence analysis and optimal matching methods in sociology: review and prospect. *Sociol Methods Res.* 2000;29:41-64.

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

- 36) Kaufman L, Rousseeuw PJ. Finding groups in data: an introduction to cluster analysis. New York: Wiley; 1990.
- 37) Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math.* 1987;20:53-65.
- 38) Hennig C. Cluster-wise assessment of cluster stability. *Comput Stat Data Anal.* 2007;52:258-71.
- 39) Hubert L, Arabie P. Comparing partitions. *J Classif.* 1985;2:193-218.
- 40) Wilson EB. Probable inference, the law of succession, and statistical inference. *J Am Stat Assoc.* 1927;22:209-12.
- 41) Tomczak M, Tomczak E. The need to report effect size estimates revisited: an overview of some recommended measures of effect size. *Trends Sport Sci.* 2014;1:19-25.
- 42) Studer M, Ritschard G, Gabadinho A, Müller NS. Discrepancy analysis of state sequences. *Sociol Methods Res.* 2011;40:471-510.
- 43) Firth D. Bias reduction of maximum likelihood estimates. *Biometrika.* 1993;80:27-38.
- 44) Heinze G, Schemper M. A solution to the problem of separation in logistic regression. *Stat Med.* 2002;21:2409-19.
- 45) Studer M, Struffolino E, Fasang AE. Estimating the relationship between time-varying covariates and trajectories: the sequence analysis multistate model procedure. *Sociol Methodol.* 2018;48:103-35.
- 46) Rojas E, Munoz-Gama J, Sepúlveda M, Capurro D. Process mining in healthcare: a literature review. *J Biomed Inform.* 2016;61:224-36.
- 47) Munoz-Gama J, Martin N, Fernandez-Llatas C, Johnson OA, Sepúlveda M, Helm E, et al. Process mining for healthcare: characteristics and challenges. *J Biomed Inform.* 2022;127:103994.
- 48) Helske S, Helske J. Mixture hidden Markov models for sequence data: the seqHMM package in R. *J Stat Softw.* 2019;88:1-32.
- 49) Murphy K, Murphy TB, Piccarreta R, Gormley IC. Clustering longitudinal life-course sequences using mixtures of exponential-distance models. *J R Stat Soc Ser A Stat Soc.* 2021;184:1414-51.
- 50) Ward JH. Hierarchical grouping to optimize an objective function. *J Am Stat Assoc.* 1963;58:236-44.
- 51) Thaddeus S, Maine D. Too far to walk: maternal mortality in context. *Soc Sci Med.* 1994;38:1091-110.
- 52) Vanasse A, Courteau M, Ethier JF. The 6W multidimensional model of care trajectories for patients with chronic ambulatory care sensitive conditions and hospital readmissions. *Public Health.* 2018;157:53-61.
- 53) Nguena Nguefack HL, Pagé MG, Choinière M, Vanasse A, Deslauriers S, Angarita-Fonseca A, et al. Distinct care trajectories among persons living with arthritic conditions: a two-year state sequence analysis. *Front Pain Res.* 2022;3:1014793. <https://doi.org/10.3389/fpain.2022.1014793>
- 54) Roux J, Grimaud O, Leray E. Use of state sequence analysis for care pathway analysis: the example of multiple sclerosis. *Stat Methods Med Res.* 2019;28:1651-63.
- 55) Kuan V, Denaxas S, Gonzalez-Izquierdo A, Direk K, Bhatti O, Husain S, et al. A chronological map of 308 physical and mental health conditions from 4 million individuals in the English National Health Service. *Lancet Digit Health.* 2019;1:e63-77.
- 56) Martin GL, Atramont A, Mazars M, Tajahmady A, Agamaliyev E, Singer M, et al. Days spent at home and mortality after critical illness: a cluster analysis using nationwide data. *Chest.* 2023;163:826-42.
- 57) Dagliati A, Malovini A, Tibollo V, Bellazzi R. A process mining pipeline to characterise COVID-19 patients' trajectories and identify relevant temporal phenotypes from EHR data. *Front Public Health.* 2022;10:815674.
- 58) Manktelow M, Iftikhar A, Bucholc M, McCann M, O'Kane M. Clinical and operational insights from data-driven care pathway mapping: a systematic review. *BMC Med Inform Decis Mak.* 2022;22:43.
- 59) Najjar A, Reinharz D, Girouard C, Gagné C. A two-step approach for mining patient treatment pathways in administrative healthcare databases. *Artif Intell Med.* 2018;87:34-48.
- 60) Adams JN, Ziegler J, McDermott M, Douglas MJ, Eber R, Gichoya JW, et al. A health equity monitoring framework based on process mining. *PLOS Digit Health.* 2024;3:e0000575.
- 61) Muyama L, Neuraz A, Coulet A. Machine learning approaches for the discovery of clinical pathways from patient data. *J Biomed Inform.* 2024;160:104746.
- 62) van der Aalst WMP. A practitioner's guide to process mining: limitations of the directly-follows graph. *Procedia Comput Sci.* 2019;164:321-8.
- 63) Guzzo A, Rullo A, Vocaturo E. Process mining applications in the healthcare domain: a comprehensive review. *WIREs Data Min Knowl Discov.* 2022;12:e1442.
- 64) Crozes F, de Boissezon X, Demeulemeester R, Mounie M, Montane E, Azema O, et al. Identifying pre- and posttraumatic brain injury pathways with a multidomain sequence analysis: a retrospective cohort study. *Ann Phys Rehabil Med.* 2026;69:102148. <https://doi.org/10.1016/j.rehab.2026.102148>

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

- 65) Bodilsen J, Dalager-Pedersen M, van de Beek D, Brouwer MC, Nielsen H. Incidence and mortality of brain abscess in Denmark: a nationwide population-based study. *Clin Microbiol Infect.* 2020;26:95-100.
- 66) Bodilsen J, D'Alessandris QG, Humphreys H, Iro MA, Klein M, Last K, et al. European Society of Clinical Microbiology and Infectious Diseases guidelines on diagnosis and treatment of brain abscess in children and adults. *Clin Microbiol Infect.* 2024;30:66-89.
- 67) Corsini Campioli C, Castillo Almeida NE, O'Horo JC, Wilson WR, Cano E, DeSimone DC, et al. Bacterial brain abscess: an outline for diagnosis and management. *Am J Med.* 2021;134:1210-7.
- 68) Muzumdar D, Jhawar S, Goel A. Brain abscess: an overview. *Int J Surg.* 2011;9:136-44.
- 69) Sharma BS, Gupta SK, Khosla VK. Current concepts in the management of pyogenic brain abscess. *Neurol India.* 2000;48:105-11.
- 70) Donabedian A. Evaluating the quality of medical care. *Milbank Mem Fund Q.* 1966;44:166-206.
- 71) Kruk ME, Gage AD, Arsenault C, Jordan K, Leslie HH, Roder-DeWan S, et al. High-quality health systems in the sustainable development goals era: time for a revolution. *Lancet Glob Health.* 2018;6:e1196-252.
- 72) Gardner EM, McLees MP, Steiner JF, Del Rio C, Burman WJ. The spectrum of engagement in HIV care and its relevance to test-and-treat strategies for prevention of HIV infection. *Clin Infect Dis.* 2011;52:793-800.
- 73) Haber N, Pillay D, Porter K, Bärnighausen T. Constructing the cascade of HIV care: methods for measurement. *Curr Opin HIV AIDS.* 2016;11:102-8.
- 74) Friedman CP, Wong AK, Blumenthal D. Achieving a nationwide learning health system. *Sci Transl Med.* 2010;2:57cm29.
- 75) Institute of Medicine. Best care at lower cost: the path to continuously learning health care in America. Washington (DC): National Academies Press; 2013.
- 76) Ovretveit J. Improving complex systems with improvement: challenges and solutions. *Front Health Serv.* 2026;5:1724893.
- 77) Punchak M, Mukhopadhyay S, Sachdev S, Hung YC, Peeters S, Rattani A, et al. Neurosurgical care: availability and access in low-income and middle-income countries. *World Neurosurg.* 2018;112:e240-54.
- 78) Fuller A, Tran T, Muhumuza M, Haglund MM. Building neurosurgical capacity in low and middle income countries. *eNeurologicalSci.* 2015;3:1-6.
- 79) Park KB, Johnson WD, Dempsey RJ. Global neurosurgery: the unmet need. *World Neurosurg.* 2016;88:32-5.
- 80) Kanmounye US, Robertson FC, Thango NS, Doe AN, Bankole NDA, Ginette PA, et al. Needs of young African neurosurgeons and residents: a cross-sectional study. *Front Surg.* 2021;8:647279.
- 81) Ukachukwu AK, Still MEH, Seas A, von Isenburg M, Fiegggen G, Malomo AO, et al. Fulfilling the specialist neurosurgical workforce needs in Africa: a systematic review and projection toward 2030. *J Neurosurg.* 2022;138:1102-13.
- 82) Barthelemy EJ, Diouf SA, Silva AS, Mukasa J, Kanmounye US, Ogunfolaji O, et al. Historical and current state of neurosurgery in sub-Saharan Africa. *World Neurosurg.* 2022;167:e670-8.
- 83) Meara JG, Leather AJM, Hagander L, Alkire BC, Alonson N, Ameh EA, et al. Global Surgery 2030: evidence and solutions for achieving health, welfare, and economic development. *Lancet.* 2015;386:569-624.
- 84) Dewan MC, Rattani A, Fiegggen G, Arraez MA, Servadei F, Boop FA, et al. Global neurosurgery: the current capacity and deficit in the provision of essential neurosurgical care. *J Neurosurg.* 2018;130:1055-64.
- 85) Rudolfson N, Dewan MC, Park KB, Shrimme MG, Meara JG, Alkire BC. The economic consequences of neurosurgical disease in low- and middle-income countries. *J Neurosurg.* 2018;130:1149-56.
- 86) Mensah EO, Mehta NH, Gal ZT, Pacheco-Barrios N, Hamzah R, Rosseau G, et al. The rising bibliometric profile of global neurosurgery. *J Clin Neurosci.* 2026;143:111725.
- 87) Rose J, Gelb AW, Davies JI, Martin J, McIntyre K, Mellin-Olsen J; Utstein Surgical Metrics Group. From definitions to implementation: a guide to collect and apply the Lancet Commission on Global Surgery indicators — an Utstein consensus report. *J Data Inf Sci.* 2026;11:46-62. <https://doi.org/10.2478/jdis-2025-0053>
- 88) Davies JI, Gelb AW, Gore-Booth J, Martin J, Mellin-Olsen J, Åkerman C, et al. Global surgery, obstetric, and anaesthesia indicator definitions and reporting: an Utstein consensus report. *PLoS Med.* 2021;18:e1003749.
- 89) Sonderman KA, Citron I, Meara JG. National surgical, obstetric and anaesthesia planning in the context of global surgery: the way forward. *World J Surg.* 2019;43:412-6.
- 90) Truché P, Shoman H, Starr N, Choa R, Ashby T, Citron I, et al. Globalization of national surgical, obstetric and anesthesia plans: the critical link between health policy and action in global surgery. *Global Health.* 2020;16:1.
- 91) World Health Assembly. Resolution WHA68.15: strengthening emergency and essential surgical care and anaesthesia as a component of universal health coverage. Geneva: World Health Organization; 2015.

Care Trajectories Rather than Outcomes: Sequence Analysis Identifies Four Hospital Pathways in Intracranial Suppurative Infections

- 92) Levesque JF, Harris MF, Russell G. Patient-centred access to health care: conceptualising access at the interface of health systems and populations. *Int J Equity Health*. 2013;12:18.
- 93) Hirschhorn LR, Baynes C, Sherr K, Chintu N, Awoonor-Williams JK, Finnegan K, et al. Approaches to ensuring and improving quality in the context of health system strengthening. *BMC Health Serv Res*. 2013;13:S8.